

nature

30 June 1977

Carry on talking

YOU MIGHT think that, what with the Select Committee on Science and Technology, the Parliamentary and Scientific Committee, The Royal Society, the Council of Engineering Institutions, the British Association for the Advancement of Science, the Council for Science and Society, Chief Scientists in governmental departments and an Advisory Board for the Research Councils, there is no shortage of channels in Britain by which scientists could make their views known to politicians and by which politicians could learn more about science and technology. You would be wrong. Scientists and politicians generally only talk to each other when things are going wrong, and even then it is usually directors of laboratories or scientific administrators who speak for science. A modest welcome, then, for the initiative of the Anglo-German Foundation for the Study of Industrial Society in trying a new format.

The foundation, in collaboration with Mr Ian Lloyd, MP, chairman of the Select Committee's sub-committee on science, recently invited a small number of MPs and staff members of the committee to spend a day and a half listening to six scientists and engineers describing their fields of interest, the prospects they see and the difficulties they face, not simply in the intellectual sphere but also in securing adequate support. The topics covered were construction,

power generation, immunology, genetics, microelectronics and the earth sciences.

In many ways the initiative was imaginative: the group was small enough for round-the-table conversation to flow easily, the scientists and engineers had ample time to describe their work, no one was angling for funds or speaking for the public record. On the other hand the MPs, for whom the show was staged, were conspicuous mostly by their absence; by half way through the gathering, Mr Lloyd was the only one left. There were, no doubt, other pressing matters to be attended to, and the harsh realities of the ballot box make little allowance for how well an MP has briefed him or herself, least of all in science or technology. So the welcome for the initiative must be tempered by some concern about whether politicians really value such exchanges. And if they do, the foundation has to decide on what lines another meeting should run. Should MPs be treated to completely new faces each time or should the effort be made, whatever the accusations of cosiness, to keep a nucleus of scientists and engineers together in the hope that good relations will develop? There is much to be said for the latter policy, tempered, of course, by new faces where necessary. For the initiative should develop into a talking shop, not a lecture series. □

Whose venture, whose gain?

THERE were signs at a recent meeting in Basel that the rising risk-consciousness reflected, for instance, in the endless debate on recombinant DNA and in the witch-hunt for environmental carcinogens, is increasingly losing the sympathy of research scientists. The meeting was not about potential environmental hazards from research; the scientists were all there at the invitation of Hoffman-La Roche to celebrate their research director's 60th birthday with a series of talks on the contribution of fundamental research to the medicines of tomorrow. Participants however seemed almost equally preoccupied with the possibility that the fruits of that research might never reach the market.

The main source of unease, of course, is the fear that far-reaching defensive action may be taken on the basis of questionable evidence. This has already happened with saccharin in the USA; whether in the end the saccharin link with bladder cancer is or is not substantiated, the fact remains that the evidence on which it has been condemned is widely regarded as inadequate. Professor Peter Cerutti implied that the over-enthusiastic use of the Ames test for carcinogens could very well lead to similar over-reactions. The Ames test depends on the correlations between mutagenic effects of chemicals on bacteria and their carcinogenic effects in whole animals. But the bacteria used in the tests are much more sensitive than animal cells, whose repair mechanisms for chromosomal damage give them complete protection from chemicals at doses that would produce a positive result in Ames bacteria.

As Professor Burgen unarguably remarked in his summing-up, "We must be cautious but not timid". After

all, nothing venture, nothing gain. But one does have to be sure that it is not the public health that is being ventured for the gain of pharmaceutical companies. And from that point of view, it is not at all clear that the public is its own best friend. A recent survey by Sir Richard Doll and his colleagues, for example, illustrates once again how enormously doctors over-prescribe anti-microbial psychotropic drugs (*British Medical Journal*, 1, 1561; 1977).

The hazards of over-use of antibiotics are well known, and the natural ability of virulent strains to acquire resistance to drugs has often been emphasised in the course of the recombinant DNA debate. The hazards of psychotropics, of which the most popular are Roche's own Librium and Valium, are not as well known. But as the authors of the *BMJ* article point out, with about 10% of males and 20% of females now taking psychotropic drugs, it is necessary to ask questions about, for instance, their possible effects on driving and other skills that require judgment.

This is not a case of the pharmaceutical industry foisting on the public drugs which carry unacceptable risks; antibiotics are life-savers, and psychotropics are certainly invaluable in some circumstances. It is the unnecessary use of the drugs that creates the unacceptable risk. However, people are not rational about the acceptability of risks. Recent publicity shows that they will insist on being allowed (or even encouraged) to smoke, or to drive without seat belts, both of which carry large known hazards, and to accept the risk of disease to an unvaccinated child rather than minuscule risk of damage by vaccination. □



Nutrition planning: the poverty of holism

Donald S. McLaren, of the Department of Physiology at Edinburgh University's Medical School, criticises the approach of the planners to the problems of world nutrition

SINCE the World Food Conference in Rome in 1974 there has been a rising tide of proposals for nutrition planning, policies and strategies on both national and international scales. The United Nations, for example, has given its blessing through the FAO/WHO report on 'Food and Nutrition Strategies in National Development', and has further encouraged the trend through the choice of 'The importance of national and international food and nutrition policies for health development' as the topic for technical discussion at the recent World Health Assembly. But the proposals are shibboleths of holism that, lamentably, have gone unchallenged.

The holistic approach has undoubtedly had its triumphs, where the problem was appropriate and the means available made it amenable. The Manhattan Project and the American and Russian space programmes are obvious examples. But the disappointing results in medicine of the American Heart, Stroke and Cancer programme and of Britain's National Health Service, and the failure of numerous attempts to control population increase, suggest that in certain areas of social action guerilla tactics and not a war machine are required, giving flexibility, accountability and proven effectiveness. Nutrition is just such an area.

The new breed of nutrition planners have a disturbing predilection for coining vague neotechnologisms and drawing maze-like flow (or are they ebb?) diagrams. That is bad enough. But they also choose to ignore those harsh realities that do not fit into their preconceived schemes. Ironically, a prime example is the phenomenon as ugly as its own 'neotechnological' name: Agripower. The probing of the United Nations Research Institute for Social Development and others, for example, recently highlighted by Susan George in *How the Other Half Dies: the Real Reasons for World Hunger*, show how the world's poor and hungry are increasingly at the mercy of vast agri-business corporations (chiefly from the United States), of Western governments with their food 'aid' policies, and of supposedly neutral multilateral development organisations, like the World Bank—all of them, so the argument runs, working in collusion with local elites nurtured and protected by the powerful in the developed world.

Then there is the case of *Scientific American* devoting an entire issue to food and agriculture. The problems of food and nutrition were there dismissed as being entirely technological in nature and capable of solution by "additional technology and capital from the developed countries"—a root cause rather than a solution. Ironically, the numerous lavish advertisements documented the frightening story of US agripower that was shunned in the text. Rarely do the institutions responsible permit themselves to make statements on the ethical implications of their activities. It probably takes someone like the former US Secretary of Agriculture, Earl Butz, to admit that "Food is a weapon. It is now one of the principle tools in our negotiating kit"—and we know what happened to him.

What has been ignored throughout such deliberations is the different nature of the two problems of food supply and human nutrition. It was this misconception which contributed to the failure of the UN Freedom from Hunger Campaign in the mid-1960s and to the Great Protein Fiasco of the early 1970s, as I pointed out in *The Lancet* at these times. Food shortage is a necessary but not adequate cause of undernutrition. There is many a slip betwixt the cup of

food supply and the lip of hungry man. Problems with the former are necessary, but not adequate, causes of those of the latter. That is why the Green Revolution has not resulted in less malnutrition.

Nutrition planners state that past efforts failed, if failed they did, because they were piecemeal and for no other reason; they claim that a new planned approach must succeed because it is holistic and not piecemeal. Now the term 'whole' has two senses, as Popper indicated in *The Poverty of Historicism*: the totality of all the properties or aspects of a thing; and the special properties of certain things which make them appear to be an organised structure rather than a 'mere heap'. Popper's argument is clear: though wholes in the latter sense have often been made the object of scientific study (notably by the 'Gestalt' school of psychology), in a social context the 'organised structure' in no way applies and holism is inappropriate. Methodological problems have not been surmounted and there is no reliable procedure for building hypotheses based on well devised observations, experiments or rational deductions. Holistic or Utopian social engineering is always public and not private, aims at remodelling the whole of society and extending the power of the state, and smacks of tyranny. The alternative is piecemeal social engineering, the word piecemeal having no pejorative sense but meaning just what it says—piece by piece or part at a time.

Past achievements of the piecemeal approach to nutritional problems, it is true, are not impressive. Iodisation of table salt is effective against simple goitre. Fortification of staple foodstuffs has helped to maintain optimum micronutrient status in many populations. Prevention of the most common deficiency, iron, remains unachieved everywhere. Classical vitamin deficiencies have been influenced most by political action and social change—the disappearance of beriberi and xerophthalmia from Japan in the early part of this century, of pellagra from the southern United States initiated by the immigration laws in the 1920s, and the emergence of rickets in British Asian immigrants. The widespread intractable problem of the malnutrition-infection complex of early childhood in developing countries has deep socio-economic roots. Mothercraft and nutritional rehabilitation centres have made modest but real contributions.

In those communities where overnutrition and nutrition-related degenerative diseases are rife, it remains to be seen whether the health of the increasingly diet-conscious public will benefit from the multifarious and often conflicting advice it is being offered. In the meantime the fact is that nutrition remains largely one of Medawar's 'kitchen arts'—a science insufficiently underpinned by theory. A halt should be called to the present mania for holistic day dreaming. Research efforts should be renewed to bring order to the present 'mere heap' so that sounder measures may result, piecemeal though they may be. The WHO's resolution at its recent assembly, urging the eradication of the severe clinical forms of malnutrition by the turn of the century, offers at least some hope. □



Hospital meals: effective against malnutrition?

To get more than enough

Colin Norman reports on a massive research effort to combat hunger and malnutrition proposed in Washington last week

A RECORD harvest in the key grain producing countries last year and the prospect of another bumper wheat crop in the United States this year have temporarily removed famine and hunger from the daily headlines. But, with international grain reserves still at a dangerously low level, a catastrophic food shortage in many parts of the world is still only one poor harvest away. Moreover, even if luck continues to hold, there's little room for complacency: according to widely accepted estimates, between 500 and 1,000 million people are already chronically underfed, and rapidly increasing population in most developing countries is continuing to drive up food demand.

Against that grim background, it may seem like the height of folly to suggest that "it should be possible to overcome the worst aspects of widespread hunger and malnutrition within one generation". But that is the conclusion of a major study published last week by the National Academy of Sciences*. One of the most intensive investigations of world food prospects yet put together, the study took two years and involved some 1,500 people; the effort was coordinated by a committee headed by Harrison Brown, professor of geochemistry at Caltech.

The committee's optimism is founded on the expectation that a massive research effort on a broad spectrum of agricultural, social and population problems could lead to a doubling of agricultural production in the developing countries and more equitable distribution of food during the next quarter century. Such an expansion would, in theory, be sufficient to cope with increased food demand from rising population and affluence, and it would also provide an increase in food consumption per capita worldwide.

But the committee is also quick to acknowledge that science and technology will not alleviate widespread hunger by themselves. Basic social changes will be needed to ensure more effective distribution of available food and to reduce poverty, though the committee's report does not advocate any specific social remedies. More attention to nutrition policies will also be required in both the poor and affluent countries, and there must also be renewed efforts to hold down the explosive rate of population

growth, the committee states.

Given the complexity of the world food problem, the committee's optimism is surprising. During the past few years there has been a fundamental shift in the food picture as grain reserves have dwindled to less than a third of their 1970 level and more countries have become dependent on the North American breadbasket to meet their food demands. After a decade of stability, crop prices have fluctuated wildly, aggravating problems of hunger and malnutrition in the poorer countries. And the increasing reliance on North American agriculture to meet world demands raises the grim possibility that a failure of the United States and Canadian harvest would have dire consequences.

Numerous and complex

The roots of the problem are numerous and complex, but one of the chief factors has been rising population in many developing countries which has caused food demand to outstrip production. According to figures in the Academy's report, although food production in the developing countries rose at an unprecedented rate of 2.8% per year between 1960 and 1975, demand for food increased by 3.5% per year. Consequently, a growing number of countries have joined the long list of food importers and there has been virtually no increase in per capita food consumption in the developing countries.

And the prospects for population growth don't give much room for optimism. Although there are recent signs that world population growth is beginning to abate, there are likely to be at least 6,000 million mouths to feed by the year 2000, and some 90% of the growth will take place in the developing countries. World population has recently passed the 4,000 million mark. To feed that many people and provide even a modest increase in per capita consumption will require Herculean efforts. According to the Academy study, the developing countries will need to increase their food production by about 3-4% per year, while stepping up efforts to reduce their birth rates and providing a more equitable distribution of food.

The task will require a major infusion of funds into research, particularly in the United States, and structural reforms in research programmes and in the political and agricultural systems of many countries,



High yield rice under test

Though the committee accepts that the task will not be easy, it states that "a latent political will now exists in numerous countries which could be mobilised in a mutually supporting fashion to commence and support such efforts".

In the past, increases in agricultural production have largely been won in two ways: increasing the acreage under cultivation and improving yields per acre. There have recently been sombre warnings, however, that marginal lands are being taken out of production in some parts of the world because of such problems as desertification and soil erosion and flooding caused by deforestation. The prospects for increasing agricultural acreage are therefore poor, and the Academy study predicts that annual increases in crop area will be unlikely to exceed 1% per year over the next 25 years. The bulk of the increase in agricultural production will have to come from boosting yields per acre.

The committee offers a shopping list of research programmes which it says should be stepped up to help increase yields. They include the development of fertilisers for use on tropical soils, efforts to improve irrigation and soil conservation, methods of reducing crop losses between the farm gate and the market, the development of more efficient farm systems and the application of biological methods of pest control. Two programmes are expected to be crucial to the long-term prospects for increasing yields, however. They involve boosting the efficiency with which crop plants use solar energy through photosynthesis, and efforts to develop cereal plants capable of utilising nitrogen directly from the atmosphere instead of having it supplied as chemical fertiliser.

At present, some 45 million tons of fertiliser is used per year worldwide. To bring about a doubling in food output during the next quarter century

*World Food and Nutrition Study, Printing and Publishing Office, National Academy of Sciences, 2101 Constitution Ave NW, Washington DC 20418, \$6.95.



A must to avoid

would require between 2 and 4 times that amount, the Academy committee reckons. That would require about 500 new fertiliser plants, at a cost of some \$50,000 million at today's prices. Such a vast use of fertilisers would not only be inordinately expensive, but it would also lead to environmental problems through run-off into streams and rivers. Hence the importance of efforts to produce crop plants which can fix their own nitrogen.

It has long been known that leguminous plants have a symbiotic relationship with bacteria living in their root systems which fix nitrogen from the atmosphere and supply it to the plant in a usable form. The immediate goal of research efforts is to improve that symbiotic relationship, thereby increasing yields of leguminous crops without increasing the application of fertilisers. In the more distant future, research efforts are aimed at inducing similar symbiotic associations between bacteria and cereal plants and, perhaps, transferring the genetic capability to fix nitrogen directly into the cells of crop plants.

Each of those goals requires fundamental biological breakthroughs, however, and they may also involve ecological problems. Though researchers are optimistic, there's no guarantee that the required breakthroughs will be forthcoming. The Academy committee recommends a major boost in funding.

The second key effort, boosting photosynthetic efficiency, also requires substantial breakthroughs. At present, most crop plants capture only about 1-3% of the solar energy they receive, but some plants, such as corn and sugarcane, do much better. Such plants seem to possess a natural mechanism for suppressing the light-aided consumption of stored foodstuffs, and the hope is that research into the underlying processes will lead to methods of increasing the photosynthetic efficiency

of other crop plants. The potential is vast: the Academy report reckons that yields of some crops could be boosted by between 50 and 100% if the research produces results. The report therefore recommends a substantial boost in funds for such research.

Although those two efforts undoubtedly provide the major hopes for a technological fix for increasing food supplies, they are both long-term, long-shot efforts. Less glamorous projects, such as improved plant breeding, soil management and pest control are likely to be more certain and yield quicker results: they should thus be pursued equally vigorously, the committee states.

Increased social research

In the long run, social and organisational changes are likely to prove just as important as technological innovations in alleviating world hunger, and thus the committee urges a sharp increase in social research. In particular, noting that countries with such diverse social and political systems as Taiwan, the People's Republic of China and Sri Lanka have managed to increase food production while lowering birth and death rates, the committee recommends a programme of comparative analysis to see whether any generalisations emerge. "Social science research on food and nutrition lags significantly behind that of biological research because of shortages of skilled personnel, suitable methods of research, data, prior research findings, research funds and effective organisation", the committee states.

Ultimately, political and social changes to combat poverty are likely to be the most effective means to reduce hunger. They are, however, less easy to achieve than changes in the technology of food production. The Academy is nevertheless relatively sanguine about the prospects. □

The Academy committee's report calls for a sharp increase in agricultural research on food production and social research on food and nutrition policies. A large proportion of the effort will probably have to be undertaken in the affluent countries, particularly in the United States, but the committee urges that steps be taken to build up the research capacities of the developing countries, and that the international network of research centres be strengthened and expanded. The chief recommendations:

- Research spending on food and nutrition problems should be increased by \$210-230 million a year in the United States. A large share of the increase should go to the Department of Agriculture, whose research budget is now about \$700 million a year. The National Science Foundation, the National Institutes of Health and the Agency for International Development should also be given large increases in their research support.

- A new programme of competitive research grants should be established in the Department of Agriculture for work on basic problems such as increasing photosynthetic efficiency and biological nitrogen fixation. (The Ford Administration in fact proposed that \$28 million be spent on such a programme in fiscal year 1978, but the House Appropriations Committee recently reduced the amount to \$10 million.)

- A new post should be established in the Department of Agriculture with responsibility for research and education.

- The United States should train more researchers from developing countries and provide more direct support for establishing facilities and research institutions in developing countries.

- Two new responsibilities should be added to the Executive Office of the President: an office to develop and maintain a coherent strategy for dealing with world food and nutrition problems, and another to coordinate US and international efforts (in fact, such tasks will probably be given to the Office of Science and Technology Policy.)

The committee briefed White House officials on the recommendations last week and, though the reception was said to be sympathetic, the discussion concerned bureaucratic problems in the Department of Agriculture rather than research strategies.



Harrison Brown, committee coordinator

Towards a pollution policy

Richard Sandbrook outlines how a UNEP plan could help to save the Mediterranean from pollution

THERE are many indicators that the so-called environmental debate has moved on from its more sensational phase into a new and more difficult stage of human endeavour, at least as far as pollution questions are concerned. The catch phrases — 'limits to growth', 'spaceship Earth', 'the environmental revolution' — do not excite the media or its readers any more, much less governments. Anti-pollution legislation is no longer pushed through the world's parliaments, the UN or the EEC on a wave of public alarm and high idealism. It is becoming increasingly necessary to present solid scientific evidence of harm and an economic evaluation of benefits and costs before controls gain political support.

This trend is becoming evident in certain sections of the United Nations Environment Programme (UNEP) too. Born in Stockholm in 1972 on a crescendo of 'environmental' enthusiasm, it is now struggling to survive in the harsh world of international law, inflation-prone economies and hard scientific evaluation. But in one area at least UNEP has made this shift successfully. Its Mediterranean Programme is one of the most complex international pollution control and environmental protection exercises ever undertaken. It reinforces the trend toward regional rather than global approaches to environmental protection — at least for the marine environment; and it helps to marshal together many of the previously competing UN agencies, governmental institutions and independent research units that can play a role.

Intergovernmental consultations on the programme started in February 1974, with the Food and Agricultural Organisation (FAO) meeting on the protection of living resources from pollution. UNEP became involved that year and in February 1975 convened an intergovernmental meeting in Barcelona on the general protection of the Mediterranean. The gathering discussed the first draft of a framework or 'umbrella' convention for the protection of the Mediterranean, and approved a three-part programme encompassing legal arrangements, scientific research and monitoring and integrated planning for the region.

A year later, in February 1976, the umbrella convention was finally approved by 16 of the 18 Mediterranean

states (the exceptions being Algeria and Albania), together with two specific protocols attached to it on the prevention of pollution by dumping (ships and aircraft) and on intergovernmental co-operation in the event of a pollution emergency. The convention was very broad, committing states to "take all appropriate measures . . . to prevent, abate and combat pollution of the Mediterranean Sea area and to protect and enhance the marine environment". A signed convention is far from a ratified convention, however, and a ratified convention is in turn far from a working one. The 1969 amendments to the IMCO 1954 oil pollution convention, for example, only came into force after the required number of ratifications in January of this year; it will be a further year at least before it is working.

Will the Mediterranean agreements suffer the same fate? One suspects not, for the convention and the two protocols to date are but a part of the much wider environmental programme — driven on, in the northern Mediterranean at least, by the threat of reduced tourism, waterborne health risks and highly contaminated marine life locally. In March 1975, for example, the French Institute of Medical Oceanography issued a warning that if mercury pollution continued at its current rate then fishermen and others who consumed on average two kilos of certain local fish a week would be chronically ill within 7 years and dead in 20 from mercury poisoning.

The legal arrangements, for example, are not going to end with ratification of the three existing instruments. A further protocol on land source pollution could be ready by the end of 1977 to be added to the framework convention, and a further protocol on exploration and exploitation of the seabed and the continental shelf is planned.

Questioning the rationale

The rationale for this grand plan, at least on an international level, is often questioned. Why not leave pollution control up to individual states or at most neighbouring states? Is it simply following a fashion set by the Oslo and London dumping conventions and the Paris convention on land source pollution? Or do the coastal states really believe that an integrated and regional approach is necessary?

The geography of the region provides some explanation. The Mediter-

anean Sea is a semi-enclosed sea and its water is effectively divided into two basins (north and south) which rotate cyclonically. In this way pollutants cross national frontiers; those from Italy potentially affect France, and those from the Nile Israel. The currents hug the coasts and such ship-source pollutants as floating litter and tar tend to be spun centrifugally toward them. Saline water enters the sea in the surface waters of the Straits of Gibraltar and leaves by the same route at greater depth. The average estimated residence time of the water is of the order of 80 years, although the range is probably from a few to several hundred years. This in effect means that for those long-lived and stable pollutants accumulation over time is bound to occur.

Secondary to the transfrontier rationale is the monitoring and research needed to underpin environmental controls. Given that the first step in controlling any pollutant is to know its source, its concentration and then its effect, the idea of monitoring the sea as a whole for the three sets of variables makes eminent good sense politically and for the more mobile and persistent pollutants good sense scientifically. This monitoring work is now progressing on seven fronts throughout the Mediterranean basin:

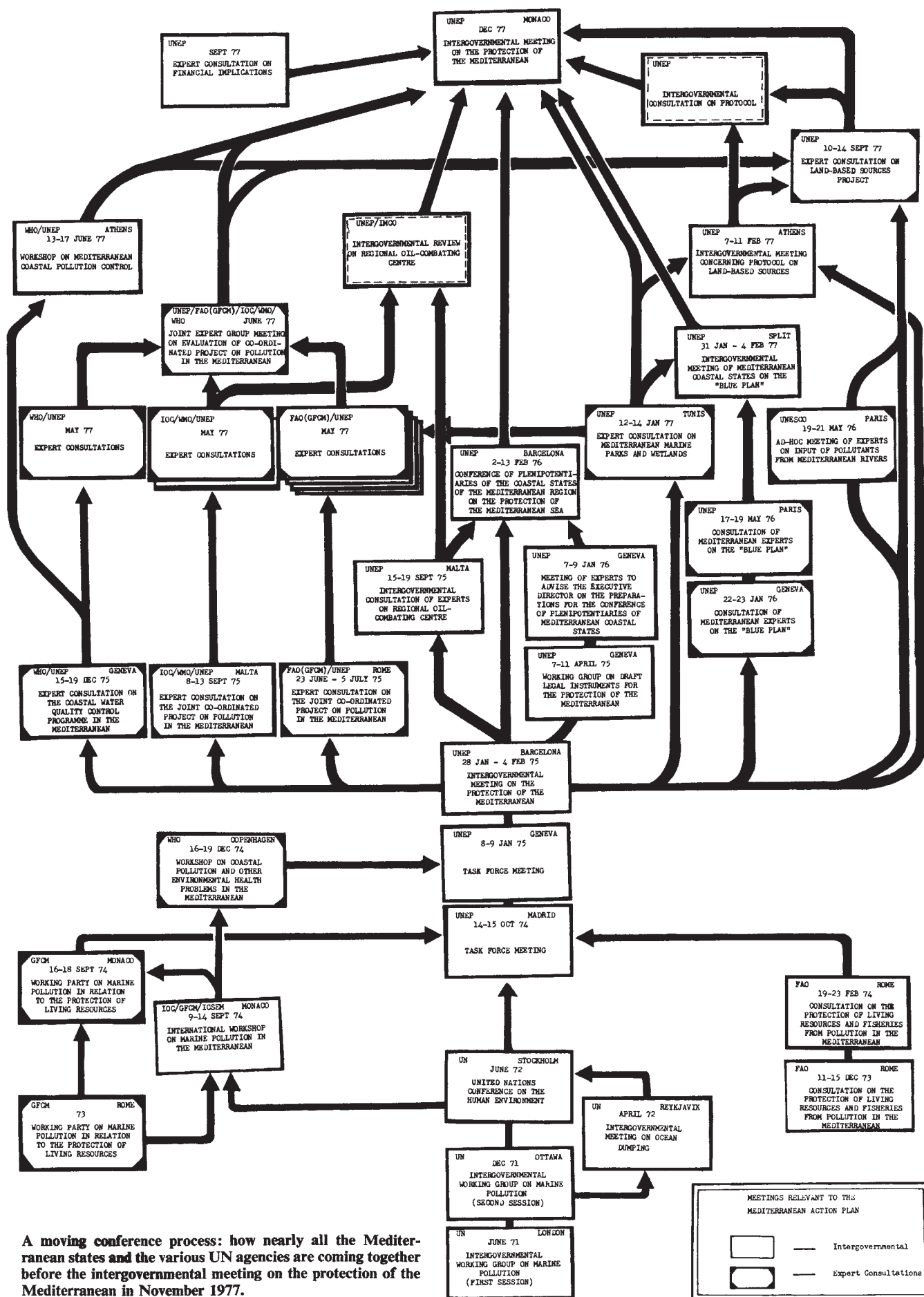
- Baseline studies and monitoring of oil and petroleum hydrocarbons, in conjunction with the Intergovernmental Oceanographic Commission of UNESCO and the World Meteorological Organisation. This programme, which now involves 11 of the 16 coastal states and 26 national laboratories, is focused on the new regional oil combating centre in Malta, which is run by IMCO.

- Baseline studies and monitoring of the heavy metals, particularly mercury and cadmium in marine organisms—a project sponsored by FAO and UNEP in 14 countries and 38 laboratories.

- Baseline studies and monitoring of DDT, PCBs and other chlorinated hydrocarbons—an FAO/UNEP project in 13 countries.

- Research on the effects of pollutants on marine organisms and their populations (FAO/UNEP in 11 countries). The aim is to study the effect of various pollutants on the population dynamics of certain test organisms; it is hoped the study will lead to the identification of particularly vulnerable species

*The 1972 Oslo Convention for the Prevention of Marine Pollution by Dumping from Ships and Aircraft entered into force in 1974 and covers the North Sea and the northwest Atlantic. The London Convention is a global convention to control the dumping of wastes at sea, and entered into force in 1975; it is now administered by IMCO. The Paris Convention for the Prevention of Marine Pollution from Land-based Sources covers the same area as the Oslo Convention and is not yet in force.



A moving conference process: how nearly all the Mediterranean states and the various UN agencies are coming together before the intergovernmental meeting on the protection of the Mediterranean in November 1977.

within the different tropic levels at various stages of their life cycles and in time lead to the identification of indicator species for water quality criteria.

●Research on the effects of pollutants on marine communities and ecosystems (FAO/UNEP in 12 countries). The project will amount to the intensive field study of various coastal ecosystems under stress, or of areas where ecosystem changes may be anticipated as a consequence of development and areas that are unpolluted and designated as marine parks for control purposes.

●Problems of coastal transport of pollutants (IOC/UNEP in 12 countries). The aim is to study the water circulation in coastal areas to understand better the physical transport of pollutants.

●Coastal water quality control (WHO/UNEP in 8 countries). The overall objective of this pilot project is to produce statistically significant data on pollution levels that may affect human health.

Each of these seven projects involves a nominated 'activity centre'; also needed is the back-up support of the IAEA and UNEP on intercalibration of equipment and analytical techniques and the provision of common maintenance services for the more sophisticated laboratory instruments. A technical training programme is under way; grants for instruments and a growing number of manuals and guidelines for the participating laboratories are available.

All of this will be the backdrop against which the next protocol will emerge. It will take a form broadly similar to that adopted in Paris in 1974 for the North East Atlantic, in Helsinki in 1974 for the Baltic Sea Area and in Brussels in 1976 for European Economic Community waters. As it is being drafted midway in the seven point scientific programme, however, the hope is that after a series of evaluation meetings of the projects involved, it will avoid some of the pitfalls that the precedents have fallen into. For example, the interim body set up to help the process of ratification of the Paris convention (IPARCOM) is the arena for much disagreement. Typical questions to be answered are: What constitutes pollution? What should be monitored? And by whom and where? For the convention simply states that by pollution is meant substances or energy introduced directly or indirectly by man that results in such deleterious effects as hazards to human health, harm to living resources and marine ecosystems, damage to amenities or interference with other legitimate uses of the sea. What is meant by 'hazards', 'harm' and 'interference' is unclear.

And on top of the work of IPARCOM (and any Mediterranean protocol developed) are the problems imposed by the European Commission directive on land source pollution (1976).

Black and grey lists

Both the Paris convention, and its later brother, the Brussels directive, include two lists (as annexes) of substances to be considered. The so-called black lists in the Paris convention relate to substances to be eliminated from the environment—for example, mercury and its compounds, persistent oils and hydrocarbons of petroleum origin, cadmium and cadmium compounds—and in the Brussels directive a similar list which is to be subject to uniform emission standards with an option for states to operate as-yet-undefined "environmental quality objectives". In both there are grey lists of substances that will be subject to less strict controls.

The treaty language is always, of course, strong enough to sound impressive environmentally—even, as some would allege, economically. One admitted purpose behind the Brussels approach is that all states should suffer the same environmental constraints no matter where the pollution occurs—an 'equalising' of competition. But in practice things are far from simple. There is little or no quantitative evaluation of acute or chronic harm to back any standards that may be set, and no recognition of the fact that receiving environments differ. For those who want to delay ratification this provides excuse enough. But a less cynical view is that unnecessary controls are not in anyone's interests other than the pollution control equipment manufacturer—and so why have them? The British have become so upset by the arguments over Paris and the Brussels directives that they have now resorted to publishing a detailed explanation of their case (DoE Pollution Paper No 11).

This time round, Mediterranean governments who have been involved in the EEC and IPARCOM process should have an idea of the likely costs involved, the necessary levels of control that should be applied and most particularly be able to avoid the problems of definition that bedevil the precedents in the field. Not least, by having some research first, many more governments will be aware of just what is under consideration.

The final concern of the programme is that of integrated planning; here the so-called Blue Plan and the Priority Action Programme are the two components. Throughout, UNEP has been at pains to stress that it is not posturing as an executive agency for the area—a sort of unofficial Brussels for the Mediterranean basin—and the Blue

Plan reflects this. There are no 'line' responsibilities envisaged but a set of 'processes' to be developed to enable governments to take the environment into account when making their own decisions. Thus systematic surveys of all major economic growth areas are called for, and so are studies of such varied topics as soil protection, water resources and urbanisation trends. The Priority Action Programme is in essence a part of the same plan but it includes areas ready for immediate action before the overall process gets underway. Again a series of cooperating national and international agencies are included.

The catch, of course, is how much notice governments will take of the environmental advice that they receive, and to what extent the whole effort will become bogged down in regional economic disparities. It all looks dangerously like the European Commission but without the teeth of its Council. To a certain extent, governments will become locked into a process that will ensure compliance—the major aid agencies (UNDP and the World Bank) that are helping some of the southern coastal states are going to be influenced by the development criteria agreed under the plan's auspices. Similarly, FAO who are developing the fisheries and WHO who are helping to build up sanitation infrastructures in the southern states are going to be closely involved. However, both components are small in resources. The first phase, costing an estimated \$1.5 million in total will, one suspects, be little more than an exchange between planners and development strategists.

This is not to knock too much. It is after all remarkable that Greece should sit with Turkey—Israel with Egypt—and talk over such important domestic strategies at all, and if the approach of the Mediterranean programme goes well it is likely to spread to other regional waters—the Persian/Arabian Gulf, the Caribbean Sea and the West African coastline for example.

There is no doubt that the unified scientific monitoring will lead to progress first in our understanding of the marine ecosystem and maybe then in controlling the obvious pollution that is going on. Governments are discovering disturbing levels of heavy metal contamination and waterborne disease in their waters. These facts are going to be made public one day and that is going to ensure a change in the level of investment and control exercised. Cousteau, who in 1972 said that the Mediterranean would surely die, admits that much progress. In 1976, he sent a telegram to UNEP saying that it might live after all. □

GENETICS

Full circle

The debate over potential hazards associated with recombinant DNA experiments has now turned full circle in the United States, with many of the people who first raised the issue in public arguing that legislation is unwarranted. Colin Norman reports

EXACTLY four years ago, a group of scientists attending the 1973 Gordon Conference on nucleic acids began to worry for the first time about the possibility that transplanting genes from disparate species into living viruses or bacteria might produce an organism with potentially hazardous properties. Discussions at that meeting led to a resolution calling for an investigation of such concerns, and the recombinant DNA debate has been picking up momentum ever since. Last week, at the Gordon Conference, some of the same scientists expressed alarm at where the debate has led.

An open letter, signed by 137 scientists at the conference—more than 75% of the attendees—warns that legislation now pending before Congress and before some state and local authorities in the United States could cripple recombinant DNA research. The letter claims, moreover, that the impetus behind the legislation comes from “exaggeration of the hypothetical hazards of recombinant DNA research that go far beyond any reasoned assessment”.

The Gordon Conferences are attended by scientists who are actively engaged in research on nucleic acids, and they thus represent people who would be directly affected by legislation. Their concerns are therefore not entirely devoid of self-interest, but they nevertheless probably reflect the views of a large proportion of the scientific community.

The call for a halt to legislation is, however, too late to have much effect and it could even be counter-productive. Committees in both the House and the Senate have now approved versions of legislation to regulate recombinant DNA research in the United States, and there is now strong and unstoppable momentum behind the legislative process. Moreover, there is a compelling argument for legislation: guidelines now governing federally-funded recombinant DNA research, issued last year by the National Institutes of Health, are not legally enforceable and do not formally apply to research supported by non-government sources.

The chief concern which led to the

formulation of the open letter at the Gordon Conference is that the legislation is likely to go well beyond simply putting the NIH guidelines into legal regulations. A bill approved by the Senate committee on human resources would establish an 11-member Presidential commission to regulate the research, for example, while a bill approved by the House health subcommittee would set up a complex procedure for approving some types of experiments on a case-by-case basis. Both versions call for new regulations to be drafted, and both would also impose stiff penalties on people caught flouting the rules.

The open letter claims that “the experience of the last four years has not given any indication of actual hazard”, and a number of scientists have recently testified before Congress that research indicates that the hazards may be much more remote than first thought. Such claims rest on evidence that there may be greater natural genetic exchange between species than once believed, that genes from higher organisms have not so far been found to be faithfully expressed in bacteria, and that bacteria bearing transplanted genes seem to be less capable of surviving in the environment than unmodified bacteria. The authors of the letter therefore suggest that while the hazards seem to be growing less plausible, controls on the research are getting more strict.

Nevertheless, few scientists are willing to argue that the research is entirely devoid of risk and thus should not be controlled. Moreover, opponents of the research have also

pointed out in Congressional testimony that evidence suggesting that the risks are less plausible is very preliminary, and that there is no justification for lessening the controls. Faced with those circumstances, Congress is not likely to drop its moves towards legislation because those most affected believe that strict legal regulations are unjustified.

The authors of the letter in fact tacitly concede that Congress is bent on legislation, and they thus urge that “should legislation . . . be deemed necessary, it ought to prescribe uniform standards throughout the country”. The idea is that Congress should prohibit state and local governments from setting their own controls on recombinant DNA research, but Congress is again unlikely to accede completely to the request. Both versions of legislation now under consideration would allow local authorities to set regulations which are at least as strict as the federal controls, but the local authorities would first have to prove that their additional restrictions are required to protect public health or the environment.

A few months ago, a number of prominent scientists reluctantly agreed that legislation is needed to ensure that controls on recombinant DNA research are applied equally to experiments funded by government or private funds. They also supported legislation on the grounds that it would stop local governments from setting their own controls. “Scientists were once urging Congress to pass legislation, and now they are telling us that legislation is unnecessary”, complained one Senate staff member last week. “That is not likely to help their case”. □

Open letter

The open letter signed by 137 scientists at the Gordon Conference reads:

“We are concerned that the benefits of recombinant DNA research will be denied to society by unnecessarily restrictive legislation.

“Four years ago, the members of the 1973 Gordon Conference on Nucleic Acids were the first to draw public attention to possible hazards of recombinant DNA research. The discussions which started at that meeting resulted in the issuance in 1976 of the NIH Guidelines for the conduct of this research.

“We, members of the 1977 Gordon Research Conference on Nucleic Acids, are now concerned that legislative measures now under consideration by Congressional, state and local authorities will set up additional regulatory machinery so unwieldy and unpredictable as to inhibit severely the further development of this field of research.

We feel that much of the stimulus for this legislative activity derives from exaggerations of the hypothetical hazards of recombinant DNA research that go far beyond any reasoned assessment.

“This meeting made apparent the dramatic emergence of new fundamental knowledge as a result of application of recombinant DNA methods. On the other hand, the experience of the last four years has not given any indication of actual hazard. Under these circumstances, an unprecedented introduction of prior restraints on scientific inquiry seems unwarranted.

“We urge that Congress consider these views. Should legislation nevertheless be deemed necessary, it ought to prescribe uniform standards throughout the country and be carefully framed so as not to impede scientific progress.

“The 137 undersigned are members of the 1977 Nucleic Acids Gordon Conference”.

SEVESO

Roche's reply

The debate continues on the Séveso disaster. Alastair Hay reports

In an effort to counter the criticism of of its subsidiary Givaudan, the Swiss multinational pharmaceutical company Hoffman-La Roche has recently given wide publicity to an internal newsletter on the environmental hazards at the Italian town of Séveso. The town was contaminated on 10 July last year by an explosive discharge from a chemical reactor manufacturing trichlorophenol. The discharge contained an extremely toxic contaminant, dioxin.

Over the past few months many press, radio and television reports have referred to the deteriorating position at Séveso. These reports quite naturally directed attention to Givaudan, the owners of the Icmesa chemical plant responsible for the contamination problem, and have been extremely critical of the company. The Roche newsletter represents an attempt on the behalf of its subsidiary to redress the balance. It claims to produce as objective an account as possible of the picture at Séveso, speaking of "considerable progress" made in the decontamination effort and the settlement of private compensation claims, and it refers to the "encouraging" health situation.

But the newsletter is primarily concerned with the prevailing position, referring only in passing to the circumstances of the actual 10 July accident. The reactor's explosive discharge is referred to as a toxic mixture of chemicals which escaped from the plant "as a result of a runaway reaction". As the cause of the accident is the subject of a legal inquiry by an Italian court, Givaudan obviously prefers not to make public pronouncements about the case.

In private, however, a company spokesman maintains that even after months of tests they still do not know what caused the accident. Givaudan can hardly deny that its Icmesa plant was the cause of the problem, but the company denies vehemently the suggestion that it was in any way negligent. "Human error" is mooted as the probable cause of the explosion.

Company officials insist that Givaudan knew of the previous accidents in trichlorophenol plants—at least nine have been publicly documented since 1949, and it is thought there may have been as many as 14. The officials say Givaudan accordingly designed the plant to incorporate safety

features to protect employees. This meant that the safety valve and vent pipe were placed to discharge the reactor contents away from the confines of the factory—with the disastrous consequences that the July accident demonstrated. The company now admits that there may have been a "lack of foresight" in anticipating this event.

In response to accusations that the company withheld information on its reaction processes, Givaudan insists that it had all the necessary permits authorising operations on the site. As for the trichlorophenol reaction, dioxin production was constantly monitored and ranged from 10–30 ppb, well below the level recorded in other plants producing trichlorophenol destined for the manufacture of the herbicide 2,4,5-T.

The company maintains that two weeks was necessary to assess the extent of the contamination in the Icmesa neighbourhood after the accident, but that as soon as reliable data was to hand, it was passed to the Italian authorities.

One of the principal problems facing the Italian authorities and Roche was the paucity of clinical information on dioxin toxicity in humans. Dr Giuseppe Reggiani, director of clinical research for Roche, says the health situation at Séveso is "reassuring". He insists, however, that certain categories of high risk groups, in particular the children with chloracne and the residents of the heavily contaminated Zone A, be constantly monitored as part of an epidemiological survey.

The Lombardy regional authority is to undertake such a survey under the auspices of its Health and Epidemiology Commission. The commission proposes to measure some 22 biochemical parameters in a population they consider likely to have been in contact with dioxin. The authority expects Givaudan to pay for the operation. The company, on the other hand, insists that before government claims can be met, due weight will have to be accorded to one of their decontamination proposals, made in August of last year. This proposal was not accepted by the Italian authorities, but would have involved the spraying of vegetation with an olive oil-cyclohexanone mixture, which would have facilitated the photoreduction of dioxin present. Givaudan claims that its programme could have resulted in the destruction of 80–90% of the dioxin.

BRITAIN

Down wind

WITH the publication last week of a 70-page report on wind energy*, the UK Department of Energy (DEN) has now completed its first round of studies on alternative energy sources. Responding to a question at a press conference last week to introduce the report, the department's chief scientist, Dr Walter Marshall, said it was "unambiguously clear" that the source which was "top of the pops" was wave power. That was followed by solar power, then geothermal and tidal power.

Final judgment, according to the head of the DEN's Energy Technology Support Unit at Harwell, Dr K. Dawson, would depend on economic and environmental considerations as well as potential contributions to total energy needs. But it was clear, he said, that wind wouldn't solve the country's energy problems. Marshall himself accorded it low priority because, he said, he has doubts whether it will be environmentally satisfactory.

That small dampener does not mean the DEN is not going to pursue the matter. The department and an industrial consortium are each pitching £75,000 into what is called a "detailed design study for a large machine installed on a prime site"; another £10,000 is coming from two interested Scottish utilities. The idea is to obtain a firmer estimate of costs, to improve the confidence with which the current cautiously optimistic estimates are held.

But, as Marshall sees it, there is a problem. If it is environmentally acceptable to use the 3,000-odd identified sites and it is possible to put aerogenerators on them, the costs might be reasonable and wind energy might seem economic even now at current prices. But, he goes on, if there is less confidence about the sites and it is left to the market to introduce wind power, the number of sites used would be small, costs would be high and wind power would hardly be economic.

Last week's report was completed before Sir Martin Ryle, the Astronomer Royal, detailed his arguments favouring wind power in an article in *Nature* on the economics of alternative energy sources. Asked whether Ryle's arguments would be incorporated into DEN work, Marshall said that, now the DEN had finished its own calculations, he was trying to get people together, including Ryle, to see what should be done next.

Marshall's personal view, that economic debate lies somewhere between resolvable disagreement on technological fact and irresolvable disagreement on religious theology, leads him to look for agreement on what should be done to resolve uncertainties. Hence his own preference, amply demonstrated in the case of wind power, for a systematic step-by-step scaling-up approach—a view which, as he points out, also goes down well with the Treasury.

Chris Sherwell

*The prospects for the generation of electricity from wind energy in the United Kingdom. Energy Paper Number 21 (HMSO, £2.25).

URANIUM

Solidarity and optimism

The Uranium Institute held its second annual meeting in London last week. Chris Sherwell reports

IT HAD been a busy year since members of the Uranium Institute last gathered in the splendid London facilities of the Institute of Electrical Engineers. The uranium cartel controversy had blown up. Anti-nuclear forces had made headway. And a new US nuclear policy had emerged. But if most of the more than 200 delegates who attended last week's meeting came somewhat chastened by the events of the previous twelve months, after two days closeted together they seemed to leave in slightly more buoyant mood.

A renewed feeling of solidarity was perhaps an inevitable outcome, but the sense of optimism that emerged from the usual business of exchanging ideas and informally discussing recent developments was a bonus. That optimism was expressed by Terry Price, the Uranium Institute's secretary general, when in summing up the formal symposium on uranium supply and demand he offered his 'scenario' for the immediate future.

Once it was appreciated that numbers mattered and there was agreement on the facts, he said, the realisation

that there are no easy answers would see the reinstatement of nuclear power to respectability. The current debate would be damped down, a large programme of thermal reactor building would go ahead and the industry's long lead-times would see a determination to push ahead on both fast breeders and reprocessing. The lasting legacies of the present period of uncertainty would be "wholly beneficial", not least the political legacy of a heightened awareness of the need to deal with proliferation.

Earlier the symposium had ranged over the most general and most specific problems relating to uranium and nuclear power. An opening paper examined projected energy requirements to the year 2000, and found that even in a 'world of internal contradictions' (rather than a 'belle époque') there was a large role for nuclear power. Later presentations discussed exploration techniques and factors affecting mining costs. Between these extremes came the most controversial issues—those of supply and demand and of proliferation—and here the real uncertainties arose.

On supply, for example, one paper on the US position contended that unless "drastic action" was taken, "there is not sufficient uranium forthcoming to

sustain a nuclear industry based upon light water reactors". That contrasts with the optimistic assessments of the Ford Foundation report which is apparently the basis of US policy. Another paper on the same subject reveals some of the findings of a US National Academy of Sciences panel; these also indicate serious supply shortages in and beyond the mid-1980s.

Supply questions easily spilled over into proliferation matters. Anthony Grey of Pancontinental Mining suggested that the US emphasis on international safeguards meant that Australia, by using its influence as a major uranium supplier, could lessen rather than increase the risk of nuclear war. Kyoji Mizumachi of the Tokyo Electric Power Company was more critical of the effects of the US anti-proliferation stance: this "drastic" change of policy had been a "great shock" to Japan. And Terry Price himself took the opportunity to criticise an erudite defence from Harvard lawyer Abram Chayes, suggesting that other countries would not alter their energy strategies without being convinced by their own advisers—and they, said Price, would be giving cautious advice. One speaker's conclusion sounded most appropriate: at the moment no quantitative conclusions about the impact of proliferation policy on the uranium market can be drawn. □

COMECON

Energy supply problem?

The 31st session of the Council of Mutual Economic Assistance (Comecon), held in Warsaw last week, reflected a growing concern of the eight member countries with raw materials and energy. Vera Rich reports

TO DATE, developments in the fields of raw materials and energy have largely been based on the 'unlimited' supplies of the Soviet Union. According to official doctrine, delays in the supply of these resources have been essentially a matter of logistics; so when the smaller Comecon members want immediate access to these resources, they have to contribute to the cost of development.

Now, however, although the speeches in Warsaw still stress the dependence of members countries on Soviet supplies, there is a tacit recognition that the Soviet Union cannot at present meet the total demands of the bloc. While acknowledging the "decisive

importance" of Soviet supplies, the Hungarian Premier, György Lazar, for example, stressed the need "to fully utilise the resources of the member countries"; he noted that Hungary was stepping up the pace of research into fresh domestic resources of fuel and raw materials. And Manea Manescu of Romania announced a massive cutback in the use of hydrocarbons for power purposes, so that they may be devoted primarily to the petrochemical industry.

The simple answer to the energy problem seems to be nuclear power. Vaclav Hula, the Czechoslovak representative to the Comecon Committee for Cooperation in Planning, put forward plans for the development of atomic power plants which would more than double the nuclear capacity of the Comecon bloc in each successive five-year plan. The final communique says the "session particularly stressed the importance of the programme for the maximum development of nuclear engineering of the Comecon countries

in providing a comprehensive solution of issues related to the development of the atomic power industry, and a solution to the entire problem of fuel and energy".

The communique recommended that an agreement on multilateral international specialisation and cooperation in the production and delivery of atomic power station equipment for 1981–1990 should be concluded next year. No formal reference was made to the potential hazards of a major nuclear programme, and the only reference to the environment was in a list for suggested inter-state conferences that would lead to "constructive business cooperation" with the capitalist countries. The Carter nuclear policy was not, apparently, even mentioned.

A major conversion to atomic power has long been a principle of Comecon energy policy, which is not hampered by ecology-minded minority pressure groups. The new resolution, therefore, is not so much a change in policy, as a recognition that it is time to proceed from theory to practice. □

IN BRIEF

Windscale contract details

In the second week of the Windscale inquiry in Britain, during the whole of which BNFL's managing director, Mr Con Allday, remained on the stand, further details of the proposed deal with Japan to reprocess spent oxide fuel emerged. Having earlier given sight of the proposed contract, BNFL published a summary of the terms detailing the amount of fuel involved (1,600 tonnes), the Japanese payment of development costs (£150 million) under the 'cost-plus' arrangements, and the value of the contract to BNFL (£400 million).

An official of the French reprocessing firm COGEMA, speaking at the Uranium Institute meeting last week, gave details of French reprocessing plans and listed the conditions attaching to the 'cost-plus' type of contract that COGEMA and BNFL have developed to fulfil "social and political requirements". These were: no financial

risk for the reprocessor in reprocessing foreign fuel; an advance from the utility of its share of the total plant investment in the form of prepayment for reprocessing services; and a contractual agreement, backed by an understanding between governments, giving the option of returning waste to the country of origin for final storage in an agreed form.

Whale quotas announced

At its Canberra Conference last week, the IWC announced the largest cut-back in the number of sperm whales which can be caught in the North Pacific. It recommended a complete moratorium on the killing of male sperm whales in the area and set the quota for females at 763; the equivalent figures adopted for this year were 4,320 and 2,800 respectively. As the North Pacific is the main whaling ground for the USSR and Japan, the two countries may find it no longer economic to

keep their whaling activities going in the area. Under the IWC's rules, however, objections to decisions can be raised up to 90 days after the conference and Japan and the USSR are expected to do this.

Although some conservationists thought the cutbacks were as favourable as could be expected, others feared that the IWC's reductions for next year were an indication of mismanagement in previous years: over-whaling had meant that stocks were now very low.

Marshall move

Dr Walter Marshall, chief scientist at the UK Department of Energy, is to leave his post and resume full time work for the UK Atomic Energy Authority. An announcement this week from the Minister, Mr Benn, says this is in view of the UKAEA's significant role in forthcoming important nuclear decisions.

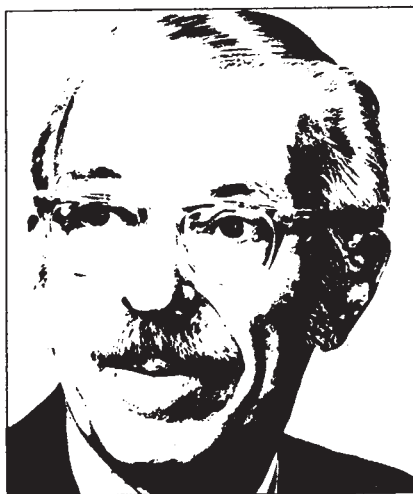
JUDGED by results, the Viking's journey to Mars last summer was an enormous success. Not the least of the wonders was successfully putting down both the landers by programmed automation, virtually blind-fold, among rocks large enough to topple them. These obstacles were too distant to be perceived in the images received on Earth before the descent started. The final 30 kilometres of drop lasted only 5 minutes. Six kilometres above the surface, the parachute blossomed out, unseen in the thin and ghostly atmosphere of Mars. Retro-rockets were ignited that pushed backwards against gravity, the lander settled gently on the surface, and its camera had taken a picture of the terrain before anyone learned of the landing. Automatic instruments chattered out the news by radio waves that took 20 minutes, at the speed of light, to get back to huge dishes, poised 350 million kilometres away in Australia, Spain and California, scooping up the incredibly faint signals.

The news was not of little green men, but of nitrogen, oxygen, krypton and xenon, of the exact temperature, and of a reddish magnetic dust on the surface. But it was the pictures of the barren, rock-strewn landscape with the metal footpad of the lander in the foreground, that spread the tidings most dramatically all over the world.

Incredibly, all the complex instruments were in working order in both landers except for one seismometer. Martian soil was scooped up and tested for life by three procedures.

The first (gas exchange) measured changes in the gas above a sample of soil moistened with a rich culture medium. The Martian soil im-

Mission to Mars



THOMAS H. JUKES

mediately released some oxygen (surprise number one). The second procedure (labelled 'release') looked for production of radioactive gas from a culture medium containing radioactive nutrients. There was rapid evolution of such gas when a fresh soil sample was added to the medium.

The third test shone a light on untreated, unmoistened soil in presence of $^{14}\text{CO}_2$ and ^{14}CO for 120 hours. The

soil was then pyrolysed and the gases that escaped were examined for products of photosynthesised organic matter. The first result was positive, although weakly so. But later experiments showed that preheating the soil to 90° for nearly 2 hours had no effect on the reaction, so the experimenters concluded "... it is unlikely the reaction is biological".

And so the initial feelings of gratification and excitement became tempered with wonder and mystification. We had steeled ourselves for negative findings (a cat without a grin), but we had not expected atypical and perhaps non-biologically positive results (a grin without a cat). The absence of the cat was underscored by negative results in tests for organic compounds in a fresh sample of soil carried out with an extraordinarily delicate 'gas chromatography mass spectrometer'.

The other instruments in the Vikings have sent a wealth of information on Martian climate, geology, meteorology and elementary composition. The North polar cap is now known to be ice. The atmospheric analysis agrees with a concept expressed by one scientist that the cap may be the tip of an iceberg; the rest of the ice is covered with Martian soil.

And so the Earth spins on its orbit delicately poised between the sterile inferno of Venus and the cold and dreary desolation of Mars. We are just the right distance from the Sun to be the planet of harvest, dew and rain; the source and abode of life, unnumbered as the sand.

correspondence

Science and ideology

SIR,—As a typically naive graduate student, I have tended to assume that science should be pursued in a rational and objective fashion and that ideology and political viewpoint have little effect on the scientific truth of a subject. Admittedly, this view stands in contradistinction to the dialectical materialism of Marxism as interpreted by Lysenko, but I have been thoroughly impressed by the advances in human knowledge that have occurred during the past several centuries using Francis Bacon's inductive approach to scientific reasoning.

At any rate, while at the recent annual meeting of the Biophysical Society in New Orleans I noticed that Dr Schwartz had organised a forum 'Racist and sexist implications of sociobiology' (note how the title is free of any ideological predisposition involving connotative and pejorative adjectives). Owing to other commitments, I missed the talk by the geneticist but I did arrive in time to catch the talk by the representative from the Boston Committee Against Racism (which I later decided should be more aptly named the Boston Committee Against Intellectual Freedom).

I had, again in my naiveté, hoped for a rational presentation of the reasons that sociobiology might be on weak scientific grounds. Instead, the talk began with the speaker admitting that he was not an expert in the area and then continuing with the unqualified assertion that the material in Edward O. Wilson's book, *Sociobiology: The New Synthesis* (1975, Harvard Press), was not scientific with the implication that no respectable scientific journal would publish such material. One of his principle objections to the theory was its use of what the speaker termed "anthropomorphism of non-human animals and insects", which, when one cut through the persiflage, meant that he strongly objected to extrapolating observations on hereditary characteristics of insects and animals to man.

It is worth noting that Lysenko used the exact same justification for outlawing medical genetics in the Soviet Union (you should see the verbal persiflage of Marxist rhetoric Lysenko used to 'prove' that human chromosomes could not exist). More mundanely,

some of my rural relatives here in Alabama use a similar rationale to 'prove' that evolution is false—God could not possibly create man along the same tree as other animals.

As the talk progressed, we were subjected to numerous emotive and pejorative terms such as racist, fascist, sexist along with the not-so-subtle hint that anyone who would view the subject from the standpoint of possible scientific validity had to be one of the above. (Reminiscent of the way that Lysenko characterised Soviet biologists who dared believe in Mendelian genetics as "enemies of the people".) The speaker certainly did not believe in a rational scientific approach but preferred to employ connotative, pejorative, obfuscatory, equivocal, circumlocutionary verbal persiflage (*sic*) that, while difficult to follow on a logical level, certainly sounded authoritative and proved even more difficult to refute. Under the circumstances, I can quite clearly see how the Eastern European IQ scare of the early 20th century is somehow related to the scientific validity of sociobiology.

Personally, I have no idea whether or not sociobiology will prove to contain elements of truth. However, I do not believe that one's ideological views are likely to change facts (even though the medieval church suppressed Aristotle's teaching for close to a thousand years because it disagreed with church doctrine). Certainly objecting to 'anthropomorphism' (the speaker's term) solely on the ground that it is objectionable flies in the face of the inductive process on which science is based. The proper approach is to make an extrapolation to man and then design experiments or search for data in an objective fashion that might prove or disprove the hypotheses involved. Unfortunately, I am not optimistic that this is likely to occur given the present political climate not only in the United States but worldwide. I am sure that any data obtained to support either the truth or falsehood of sociobiology, IQ testing or any of the related areas will be denounced by the other side—or in many cases buried under numbers and statistics from which one can, through convolution, draw one's own interpretation of the data (see *Matters Arising, Nature*, 266, 279–281; 1977).

Certainly the fact that the Bio-

physical Society would invite a speaker who was making such an overtly ideological attack speaks poorly for the organisers of the conference. I prefer to listen to my relatives in the rural Deep South discuss how evolution is a Godless bolshevik conspiracy and how Darwin was one of Satan's emissaries. Their tone is admittedly anti-intellectual, but at least it isn't ostentatiously pretentious. If one removes some of the intellectual facade from many of the ersatz scientific debates heard in the areas discussed in this letter, one will find the intellectual content comparable.

Yours faithfully,

WILLIAM O. ROMINE

*The University of Alabama
in Birmingham,
Alabama, USA*

Israel's President

SIR,—Contrary to the belief expressed (9 June, page 486), the office of President of Israel has not always been occupied by a scientist.

The first President, Dr Chaim Weizmann, was a biochemist. He was the first reader in biological chemistry at the University of Manchester and later became Director of the David Seiff Institute at Rehovot—subsequently renamed in his honour.

The second and third Presidents, Itzhak ben Zvi and Zalman Shazar, were primarily both politicians of the establishment, although ben Zvi did do some sociological research on ethnic minorities among the Jews. Before becoming President, ben Zvi was known as a founder, with David ben Gurion, of the Histadrut—the trade union movement. Shazar for many years was a newspaper editor before entering politics and subsequently becoming Minister of Education.

Ephraim Katzir, the fourth and current President, is a biochemist of international repute. He is fortunate that the duties of office as President also enable him not to neglect his scientific career.

Perhaps the message which should be passed on is 'Biochemists Rule . . . OK!'

Yours faithfully,

B. SAMPSON

*Federation of Zionist Youth,
London, UK*

news and views

Contractile filament organisation, mechanics . . .

from John Squire

FEW people would question nowadays that most muscle types, including vertebrate smooth muscle, shorten by a mechanism involving the relative sliding of two principal muscle filaments, the thin actin filaments and the thick myosin filaments. The surfaces of the myosin filaments are covered with arrays of globular projections (the ATP-splitting heads of the component myosin molecules) and the relative filament sliding is thought to be due to the interaction of these projections (often called crossbridges) with the neighbouring actin filaments. It has been shown that virtually identical actin filaments (except for differences in length and the presence or absence of the regulatory protein troponin) occur in all muscle types and also that the thick filaments, although far from being identical, probably have their myosin projections or crossbridges organised into closely related surface arrays. An inevitable conclusion from these results is that the local cross-bridge environments in all muscles are probably rather similar and that the basic force-generating mechanisms must therefore be similar too. Much current research on muscle is therefore aimed at answering two fundamental questions: 'how does the crossbridge interaction with actin actually generate force?' and, less obvious but equally crucial, 'what is the minimum organisation of actin and myosin molecules sufficient to form a viable contractile system?'. This latter question has become even more important in the light of recent fascinating discoveries of the presence of actin and myosin molecules in a bewildering variety of non-muscle cells (see Pollard and Weihing *CRC Critical Rev. Biochem.* 2, 1; 1974). It now seems evident that many, if not all cells possess the ability to form rudimentary contractile systems involving actin and myosin and possibly that these systems comprise filaments which are formed spontaneously at different places and times as required by the immediate needs of the cell.

From studies of the molecular

arrangements of the ordered arrays of actin and myosin filaments in different muscles it has become apparent that a common organisational feature crucial to muscular contraction concerns the relative polarities of the interacting filaments. Although different muscles may involve slightly different polarity arrangements the structure of a sarcomere in striated muscle (Fig. 1) serves to illustrate the point. Here the myosin filaments are bipolar; the myosin molecules of each half of a filament point in opposite directions. The actin filaments are arranged so that those which interact with one end of a thick filament are of opposite polarity to those interacting with the other end.

Organisation

The organisation of filament polarity has been illuminated by a number of *in vitro* studies. For example it was shown by Huxley (*J. molec. Biol.* 7, 281; 1963) that synthetic self-assembled myosin filaments with a bipolar structure can be prepared from relatively pure preparations of myosin. Bipolar assembly must therefore be an intrinsic property of the myosin molecules themselves. Similarly, self-assembled actin filaments can be prepared from actin solutions and these have been shown to be polar by labelling them with a preparation of isolated myosin heads (myosin heads which have been cleaved enzymatically from the rest of the myosin molecule). Such 'decorated' actin filaments show a characteristic 'arrowhead' polar structure, (Huxley *op. cit.*; Moore *et al. J. molec. Biol.* 50, 279; 1970) which shows clearly that

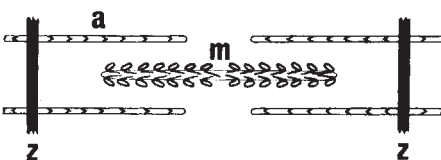


Fig. 1 Highly schematic diagram of the organisation of polarities in a striated muscle sarcomere. The Z-bands (Z) define the ends of the sarcomere. (m is a myosin filament, a an actin filament). Polarity in the thin filaments is indicated by the arrow heads.

the two molecules need to be pointing in the right directions for their highly specific interaction to occur. Polarity organisation must therefore be a crucial requirement for a viable contractile system.

In a highly ordered muscle, such as a vertebrate skeletal muscle, one could speculate that the assembly of myofibrils might well be organised systematically during myogenesis by external factors so that, for example, actin filaments of the correct polarity are laid down at each end of the sarcomere. But it is not so clear that a comparable external mechanism could be involved in organising the filament polarities in spontaneously formed contractile assemblies in non-muscle cells.

Fortunately, in a recent study of actin/myosin self-assembly, evidence has been obtained which reduces the need to postulate an external factor in the assembly process. Hayashi *et al.* (*J. molec. Biol.* 111, 159; 1977) have described the results of experiments in which globular actin molecules were added to a solution of synthetic myosin filaments to give a molar ratio of 2:1 and conditions were then altered so that the actin molecules polymerised into filaments. The structures which resulted were filament assemblies comprising bipolar thick filaments interacting at each end with thin filaments of opposite polarity; the latter presumably being defined by the myosin crossbridge polarity. These self-assembled contractile units are in many ways analogous to a small part of a muscle sarcomere with roughly six thin filaments interacting with each end of a thick filament. The only thing obviously lacking is a structure analogous to the Z-band to link the contractile units end-to-end.

Myogenesis is much more complicated, since there are many additional proteins in the sarcomere and, among other things, the thick and thin filaments have very precisely defined lengths. But these new results may show how spontaneous formation of relatively simple contractile systems might occur in non-muscle cells. Whatever the specific polarity organisation in the thick filaments this organisation may itself serve as a template for

ordering the polarities of the thin filament arrays. Since some of these non-muscle cells contain α -actinin, a protein also associated with muscle Z-bands, and presumably all the necessary regulatory proteins as well, it seems that most of the ingredients needed for the self-assembly of a simple contractile system may already be known. But one thing which is not clear from the new results is the actual mode of formation of the bipolar contractile units. The possibility favoured by Hayashi *et al.* involves the binding of actin monomers to myosin cross-bridges before actin polymerisation occurs. But the alternative, that actin polymerises and the resulting filaments then interact with the appropriately oriented end of a myosin filament, cannot be ruled out. A useful follow-up to this study would be to decide clearly between these alternatives possibly by determining the state of aggregation of the actin and myosin before polymerisation of the actin.

Crossbridge interaction

Turning now to the nature of the crossbridge interaction itself, equatorial X-ray diffraction evidence from vertebrate skeletal muscles and insect flight muscles has indicated that on activation there is a movement of mass, presumed to be the myosin cross-bridges, towards the actin filaments. Although attachment to actin probably occurs this has not actually yet been proved. Huxley (*J. molec. Biol.* **37**, 507; 1968) took the movement to be largely radial from the myosin to the actin filaments. On this basis the amount of mass transfer in active vertebrate skeletal muscle, calibrated against the transfer in rigor muscle in which actin labelling by 50 to 75% of the available myosin heads probably occurs, is about 50% of the rigor transfer (Haselgrove & Huxley *J. molec. Biol.* **77**, 549; 1973) or up to 80% (Matsubara *et al.* *Nature* **255**, 728; 1975). Cardiac muscle must be the most specialised of all the vertebrate striated muscles; it contracts rhythmically and is not voluntarily controlled. But a recent equatorial X-ray diffraction study of contracting heart muscle (Matsubara *et al.* *J. molec. Biol.* **111**, 121; 1977) has shown that a mass transfer corresponding to that in skeletal muscle occurs in this muscle too. Assuming the crossbridge movement to be radial, about 20% transfer (compared with rigor muscle) occurs during the contracting systolic phase and even during the diastolic phase there seem to be more cross-bridges near to the thin filaments (7 to 9% of the rigor transfer) than in a truly quiescent muscle. Sub-maximal tension generation was said to account for the small systolic transfer (20%)

relative to the 50 to 80% in skeletal muscle. Despite this it seems clear that comparable crossbridge movements occur in all these different muscles. But it remains to be determined what these movements really are. Are they principally radial as suggested by Huxley (see also Haselgrove *et al.* *Nature*,

261, 606; 1976) or would azimuthal crossbridge movements account more satisfactorily for the observations (Lymn & Cohen *Nature*, **258**, 770; 1975)? No doubt this fundamental problem of interpretation will be central to future equatorial X-ray diffraction studies of muscle. □

... and biochemistry

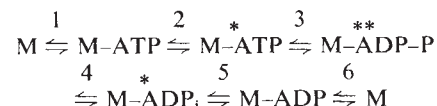
from Howard White

UNDERSTANDING the way in which muscle and other contractile systems couple the production of mechanical work to the hydrolysis of ATP will almost certainly require a detailed understanding of the mechanism of ATP hydrolysis by myosin and actomyosin (the contractile system composed of actin and myosin) and a correlation of the biochemical mechanism with the physical events of the contractile process. It is more or less accepted as dogma that during contraction the heads of the myosin molecules form cross bridges with the actin filaments, which attach and detach in a 'rowing' motion.

The mechanisms of myosin and actomyosin ATP hydrolysis (equation 1) that were outlined six years ago by Lymn and Taylor (*Biochemistry* **10**, 4617; 1971) have been especially useful because they correlated biochemistry with the cross-bridge cycle in a natural way. They showed that ATP binds to actomyosin and dissociates the myosin heads from actin before the ATP is hydrolysed. The overall rate of myosin ATP hydrolysis in the presence of actin and in contracting muscle is several hundred times more rapid than myosin alone because the reattachment of the myosin heads to actin increases the rate of the slow step of the mechanism—product release. The drive stroke in muscle presumably occurs with the return of the cross bridges to the original attached state of actomyosin without bound nucleotide.

Since Lymn and Taylor's work, the mechanism of ATP hydrolysis by myosin (or rather proteolytic fragments of myosin, S1 and HMM, which are soluble at physiological conditions and more tractable for solution chemistry) has been studied in detail. The evidence which shows that there are at least five intermediates involved in the reaction has been recently reviewed by Trentham (*Q. Rev. Biophys.* **9**, 217; 1976).

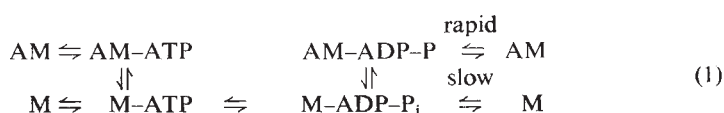
Both ATP (Step 1 and 2) and ADP (Step 6 and 7) bind to myosin in two-step reactions. The initial weakly bound



nucleotide complexes, M-ATP and M-ADP are converted at apparently identical rates to higher affinity complexes M*-ATP and M*-ADP. The transition of the high affinity complexes is associated with changes in the ultraviolet absorption and fluorescence (the number of stars given to a state gives a rough indication of the degree of enhancement of the tryptophan fluorescence). The major difference between the myosin ATP and ADP complexes is that the rate constant for M*-AXP $\rightarrow$ M-AXP is much smaller for ATP than ADP and hence ATP binds with a much higher affinity ($K_1K_2 \approx 10^{11} M^{-1}$), than ADP ($K_5^{-1}K_6^{-1} \approx 10^6 M^{-1}$). The hydrolysis of ATP bound to myosin (Step 3) is rapid and surprisingly reversible, $K_3 = 10$. The rate-limiting step of myosin ATP hydrolysis is associated with the release of phosphate from the myosin thus making M*-ADP-P_i the predominant steady state intermediate. The favourable free energy of hydrolysis of ATP in solution is therefore expressed primarily in the difference between the binding energies of ATP and ADP.

This type of kinetic approach is now being used to determine how actin modifies the rate and equilibrium constants of the myosin-catalysed ATP hydrolysis, and in particular to identify the rate-limiting step, and show which intermediates are common for both the myosin and actomyosin mechanisms.

Chock, Chock and Eisenberg, (*Biochemistry* **15**, 3244; 1976), and Sleep and Taylor (*Biochemistry* **15**, 5413; 1976) report that the first intermediate after dissociation of myosin-S1 from actin by ATP undergoes a fluorescence enhancement that is similar to that observed for ATP binding to myosin. Both laboratories mixed the actomyosin-S1 with a high concentration of ATP to ensure that the



change in light scattering due to dissociation of the actomyosin would become complete within the mixing times of their stopped-flow equipment and not complicate the interpretation of the fluorescence change. Although results and interpretation of the two groups differ slightly in detail, both reported that the change in tryptophan fluorescence which occurs after the dissociation of myosin from actin has a similar rate and amplitude to the fluorescence change observed for ATP binding to myosin. Sleep and Taylor were able to show by rapid quenching experiments that the ATP bound to the intermediate did not exchange with the ATP in solution (as would be expected for M-ATP). They dissociated actomyosin with stoichiometric ^{32}P -ATP and found that an excess of cold ATP added before the fluorescence increase had occurred did not reduce the amount of ^{32}P -ATP hydrolysed. It therefore seems most likely that the intermediate first released from actin is $\text{M}^*\text{-ATP}$ and that the fluorescence change occurs during the transition $\text{M}^*\text{-ATP} \rightarrow \text{M}^{**}\text{-ADP-P}_i$.

These observations are especially important, first because they link the change in optical properties of nucleotide binding to myosin with the change in optical properties of nucleotide binding to actomyosin and second, because they show that a change in fluorescence occurs at a step of the reaction that could be associated with the movement of the unattached myosin crossbridge. However, caution is necessary before suggesting that the fluorescence change is an indicator of crossbridge movement. The changes in optical properties observed for myosin and actomyosin are not unusual for the binding of large ligands to protein molecules. Furthermore, there is a conspicuous absence from the muscle literature of a demonstration that ATP induces gross physical changes in the structure of myosin, myosin filaments or myosin proteolytic fragments.

Chock, Chock and Eisenberg (*Biochemistry op. cit.*; 1976) (and for that matter Taylor and myself (*Biochemistry* 15, 5818; 1976) have found that the rate of recombination of myosin-S1-ADP-P_i with low concentration of actin, measured by turbidity, is proportional to actin concentration and equal to the steady state rate of ATP hydrolysis; this would be expected from the Lymn-Taylor model at actin concentrations where the rate limiting step is recombination of M-ADP-P_i to actin. Eisenberg also found that at rather low ionic strength and temperature ($\mu = 10 \text{ mM}$, 4°C) the steady state rate of ATP hydrolysis and recombination of S1-ADP-P_i and actin both reached a maximum of 1.0 s^{-1} at higher actin concentrations, and that the actin and

myosin remain mostly dissociated even at the highest actin concentrations. If a step associated with ADP or phosphate release were rate limiting for actomyosin ATP hydrolysis at high actin concentration, then one would expect that the predominant intermediate would be actomyosin-S1-ADP-P_i; that is, most of the myosin-S1 would be bound to actin. Because the covalent ATP bond splitting was thought to be rapid ($\sim 100 \text{ s}^{-1}$ at 20°C), they reasoned that there must be another step in the reaction, a rate-limiting transition from a myosin-S1-ADP-P_i intermediate unable to bind to actin (the 'refractory state') to an intermediate capable of interacting with actin. However, Taylor (*Biochemistry* 16, 432; 1977) has extended the original measurements of the rate of the covalent ATP bond hydrolysis and found it to be exceedingly temperature dependent. At 4°C the ATP is cleaved at only 3 s^{-1} , a value quite near the maximum rate of ATP hydrolysis, 1.0 s^{-1} , measured for actin-S1 by Eisenberg at the same temperature. It is therefore possible that the rate-limiting step could be the hydrolytic step and the refractory state could be $\text{M}^*\text{-ATP}$. Moreover, it is unclear whether the rate limiting step of actomyosin ATP hydrolysis is like myosin in being associated with product release or whether it is the bond-splitting step.

One may quite reasonably wonder how

the rate and equilibrium constants determined in solution can be applied to the highly organised contractile apparatus in muscle. Taylor and I have calculated the "effective standard free energy changes" between each of the intermediates of actomyosin ATP hydrolysis using the equilibrium constants measured in solution at nearly physiological conditions and using as the standard state the concentrations of ATP, ADP, and phosphate normally present in muscle. We found that more than 50% of the "effective standard free energy change" during ATP hydrolysis is associated with the release of ADP and phosphate from actomyosin. Intuitively, it would be expected that the steps with the largest effective free energy changes in solution ought to be the ones coupled to the production of work *in vivo*. However, as has been pointed out in a very elegant theoretical analysis by Hill and Simmons (*Proc. natn. Acad. Sci. U.S.A.* 73, 36, 95; 1976) and Hill and Eisenberg (*Biochemistry* 15, 1629; 1976) restriction of actin and myosin to a given three-dimensional structure could drastically alter the distribution of the equilibrium constants between the biochemical states.

The next stage of this type of work might be to measure the rate and equilibrium constants in myofibrils or muscle fibres where the three dimensional structure and ability to do work is maintained.

New technique for surface crystallography

from John Pendry

THE detailed knowledge of atomic arrangement provided by X-ray and electron diffraction is the cornerstone of our understanding of solids. These techniques have been enormously successful and almost all commonly studied crystals, at least in the area of the physical sciences, have had their atomic coordinates determined. Recently much interest has been focused on surfaces of solids, especially on the adsorption of foreign atoms, an interest driven not only by the problems of understanding these systems, but also by the obvious relevance to more applied areas. The same basic requirements for surface atomic geometry exist as for bulk but the absence of a simple effective technique applicable to surfaces is an impediment to progress. Diffraction of X rays cannot be used because the weak signal from a single layer of atoms on a surface would be swamped by the bulk signal, and the same is true for high energy electron diffraction. By making use of the increased

scattering cross sections of atoms for low energy electrons (0-200 eV) it is possible to do experiments sensitive to the surface and many simple surface structures have been solved in this way, but the experiments are not easy to perform. At the same time theoretical interpretation is involved, limiting accuracy to the $0.1 \text{ \AA}$ level, and effectively blocks extension of the method to complex surface structures. The search is still on for the definitive surface crystallographic technique.

In recent experiments performed on the Stanford synchrotron radiation project (reported by P. H. Citrin at the March 1977 meeting of the American Physical Society in San Diego; see also Stern *et al.* *Phys. Rev. Lett.* 38, 767; 1977), a revolutionary way of solving surface structures has been demonstrated. In one recent experiment the adsorption of iodine atoms on to a face of a single crystal of silver was studied. X rays were used but not in a diffracting mode: instead their energy was tuned close to one of the inner-electron energies of an iodine atom. The ejected electron behaves like a radar pulse originating from the excited atom. Surrounding atoms can scatter the out-

Howard White is Helen Hay Whitney Foundation Fellow at the MRC Cell Biophysics Unit, Kings College, London.

John Pendry is Head of the Theoretical Physics Group at the Daresbury Laboratory.

going electron back to the source where the scattered electron wave interferes with the outgoing wave, modulating the probability of electron emission relative to the free atom value. Figure 1 shows a sketch of the process. The following formula describes the modulation to a good approximation:

$$\chi(k) = \frac{I(\text{surface})}{I(\text{free atom})}$$

$$= \sum_j f_j \frac{\sin(2kR_j + \psi_j + 2\delta)}{kR_j^2}$$

where

$$\frac{\hbar^2 k^2}{2m_e} = \hbar\omega - E_c$$

E_c is the binding energy of the core level, $\hbar\omega$ is the photon energy, R_j the distance to the j th neighbour, f_j and ψ_j the amplitude and phase of scattering by the j th neighbour, and 2δ a shift in phase of the wave, caused by refraction in the potential of the emitting atom. The structure is commonly referred to as EXAFS, standing for extended X-ray absorption fine structure, but it must be stressed that the structure is in fact generated by the diffraction of electrons produced by X rays within the surface. χ can be measured for a whole range of k values by increasing the energy of the input photon, so that χ can be Fourier transformed to pick out the various sine-wave components and hence to find R_j . Some corrections must be applied for the k -dependence of f_j , ψ_j and δ , but these can be calculated quite accurately from theory.

The question remains of how to measure $\chi(k)$. X rays of variable energy can be obtained by monochromating a continuum source. High power X-ray tubes have been used, but by far the best source is a synchrotron or storage ring

which emits very powerfully in the X-ray region. The innermost shell, the K shell, is usually chosen for the experiment and cross sections are sufficiently large for a substantial number of surface atoms to be excited. The probability of electron emission can then be monitored in one of several ways. When the technique is applied to bulk atoms the electron emission probability can most simply be monitored by the absorption coefficient for X rays, but for a surface experiment this method is not sufficiently sensitive. Alternatively the presence of excited atoms can be detected by their eventual decay by way of an Auger process involving emission of a second electron of fixed energy independent of the first electron's emission energy. Sensitive electron detectors are available and the Auger signals can easily be detected. It was in this way that χ was measured for the iodine on silver investigation.

The experiments have shown that it is possible to use the EXAFS technique at surfaces and have confirmed an earlier determination of the structure by electron diffraction. Interpretation of these experiments is sufficiently simple and well understood that ultimately an accuracy of 0.01 Å should be obtainable, an order of magnitude better than previously possible, and in addition extension to more complex structures is feasible. There are limitations to the technique. Only signals from surface atoms of a type different from those in the bulk can be examined. Not all atoms have inner levels suitable for the experiment because the current generation of X-ray monochromators cuts off at long wavelengths leaving the light elements such as oxygen and carbon inaccessible for the moment. Even if an inner shell can be excited, a suitable Auger transition for the monitoring process must be found. Despite these limitations surface EXAFS is the most striking advance in surface crystallography in recent years. □



appears to be due, says Liane Reif-Lehrer of Harvard Medical School, to monosodium glutamate, which is also used in Japanese, Italian and French restaurants. All-American establishments seem to be blameless.

Reif-Lehrer brought the full resources of modern investigative techniques to bear on the problem, and the encoded results were analysed on a computer at the Eye Research Institute of the Retina Foundation. Over 2,000 individuals of all ages were interrogated, including groups of grade-school children in the towns of Needham and Wellesley, Massachusetts; "the socioeconomic situation of these towns was such that these children were likely to have had multiple exposure to Chinese restaurant food". In addition, public notices were issued encouraging sufferers to describe their experiences. To test whether those who exhibited symptoms were more likely to reply to the questionnaire than those who had never suffered, and who therefore put it in the wastebasket, the non-responders were further pursued by post and by telephone; nevertheless an obstinate 3 still refused to answer the questions. A further 22 "did not answer telephone or had terminated appointment". How many of these had suffered nervous breakdowns on account of all the pestering is not recorded.

Among those who gladly volunteered to describe their symptoms, the majority of those susceptible to CRS (about 25% of the total) reported combinations of such commonplace but disagreeable symptoms as headaches and vomiting, diarrhoea and dizziness, sneezing and stomach cramps; not to mention tightness around the face and burning in the upper torso. It is, however, in the list of 36 "other symptoms" that the full horrors of CRS emerge. Depression, hallucinations and an urge to laugh and cry were reported; also excessive perspiration, a feeling of having heavy arms and legs, or "head feels unattached". Less alarm-

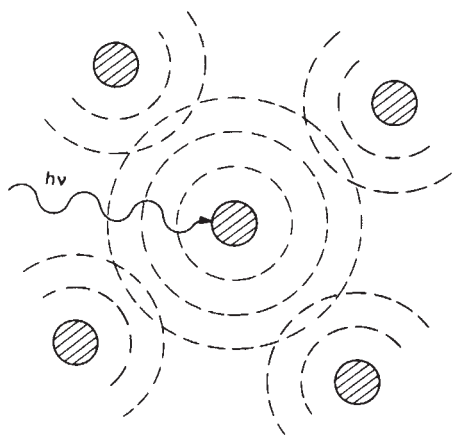


Fig. 1 Electrons emitted by an activated atom are bounced back off neighbouring atoms with a phase lag determined by the interatomic distance and the electron energy.

Advances in Chinese Restaurant research

from a Correspondent

INVESTIGATION of the dreaded Chinese Restaurant Syndrome (CRS) continues to grip the Boston medical research community like a fever. Another comprehensive account of the range of symptoms induced in some of those who (despite it all) still eat Chinese food has just appeared in *Federation Proceedings* (36, 1617; 1977). The list bears comparison with any mediaeval catalogue of bodily ailments. It all

ing were feelings of drowsiness or a "sense of fullness after limited amount of food". Four respondents blamed their Chinese meal for insomnia, while one said it made him feel "high". Another complained of loss of reflexes, making it hard to carry out typing. Others felt paralysed in the neck or backbone, and one unfortunate had a "mental image of a buzzing sound". There was also a wide range of ocular disturbances, from blurred vision to scintillating scotoma (the research was supported by a grant from the National Eye Institute). The most dramatic description, sparing no detail, was "Burst blood vessel and occasional sensation as if pins were being stuck in the eyeball".

All these fascinating data were analysed exhaustively with respect to the respondents' age and sex, though 337 people unaccountably failed to disclose the latter. It seemed that the 26-30 age group was the most at risk. Replies from the 317 children tested were dis-

missed for some reason as "far less reliable than the adult data", but their descriptions have nevertheless the ring of truth; many of those afflicted reported their only reaction as being a stomach ache, while some felt "fat", "stuffed", or "very full" after a Chinese meal. Two children did say that they "felt very funny" and others reported milder reactions, but none described such an interesting variety of bizarre symptoms as their elders, most of whom were faculty members or college students. Perhaps the children's vocabularies were insufficiently developed.

The author ends with thanks to a veritable army of helpers, including five who "stuffed envelopes". Her appetite for symptoms seems only whetted by this wide-ranging exercise, and she says "we hope to provide more definitive data on this matter in the near future". The First International CRS Symposium can hardly be far away. □

Phobos and Deimos, the moons of Mars

from David W. Hughes

A discussion meeting to review our present knowledge of the Martian satellites was held by the Royal Astronomical Society in London on 13 May, 1977, under the chairmanship of Dr G. E. Hunt of the Meteorological Office.

AUGUST 12, 1977 is the hundredth anniversary of the discovery of Phobos and Deimos and meetings are being held in London and Washington to commemorate this event. The discovery was made by Asaph Hall, an astronomer at the US Naval Observatory, Washington DC, using their new 26-inch refractor at the old observatory down by the Potomac river. In 1877 this was one of the best telescopes in the world. Hall started his search in early August, his first few nights being unsuccessful. Discouraged, he was all for giving up but his wife (*née* Angeline Stickney) urged him to spend a few more nights at the telescope and on the 12th he discovered Deimos. While confirming this observation Hall discovered Phobos on the 17 August, the two new satellites being announced on 19 August. Hall named them Phobos (fear) and Deimos (terror or panic) after the horses that pulled the chariot of the god of war in Greek

mythology. Interestingly the two Martian moons had been discussed in the literature before their discovery. Johannes Kepler (1571-1630), a great lover of mathematical progressions, had decided that as Earth had one satellite and Jupiter four (only the Galilean satellites were known at that time) then Mars, the planet in between, should have two. Voltaire and Swift mentioned them in *Micromegas* and *Gulliver's Travels*, Swift in 1726 going so far as to predict the correct values for their orbital revolution periods.

Hall observed Mars during a very favourable opposition when the distance between Earth and Mars was at a minimum (~ 0.57 a.u.). At their greatest elongation from Mars the two satellites are only 22 and 56 arc s from the centre. Even here it is extremely difficult to pick them out from the halo of scattered light that surrounds the telescopic image of Mars. The fact that they are about 3×10^8 times fainter than Mars does not help much either.

At the London meeting D. Pascu (US Naval Observatory) discussed Earth-based observations of the satellites with special reference to the accidental and systematic errors inherent in positional observations. Before 1941 visual measurements were made of the distance between the satellites and of the angle the two subtended at the centre of the planetary disk. The rapid motion of the satellites and the problem of estimating the disk centre, added to

Rat readings

If to have a newsletter devoted to one confers the ultimate cachet of respectability in the animal world, the laboratory rat has been unaccountably neglected. A new periodical 'Rat News Letter', published by the UK Medical Research Council's Laboratory Animal Centre now puts the rat on a par with several of its laboratory companions. The purpose of the newsletter, which is intended to appear twice a year, is to serve as a medium for the informal exchange of information on all aspects of the biology of the laboratory rat, but with special emphasis on the use of genetically defined rats in biomedical research.

The first issue announces the discovery of an athymic nude mutation which occurred spontaneously in the hooded rat colony at the Rowett Research Institute. This mutant is being bred at the Laboratory Animal Centre and will probably begin to become available in small numbers by the end of the year.

The vexed question of nomenclature for the alloantigenic loci is discussed and the issue also contains contributions from research establishments throughout the world giving details of their interests, and holding of strains, stocks and mutants. The editor is Dr Michael Festing, MRC Laboratory Animal Centre, Woodmansterne Road, Carshalton, Surrey SM5 4EF to whom contributions for publication and enquiries about subscriptions (£5.00/\$10.00 for four issues) should be sent.

problems of visual astigmatism, made these measurements very difficult. Photography has been used since 1941, small circular or lozenge shaped nicrome filters deposited on optically flat glass plates being used to equate the brightness of the planet and satellite. This technique obviously reduces the personal error and with exposure times mainly between 10 and 20 s the satellites do not move very far during the exposure time. Photography thus reduces the positional error from about 0.6 arc s to less than 0.1 arc s.

Photometric observations are greatly hampered by the logarithmic intensity gradient of the planetary halo and only integrated results over the whole satellite disk are possible and these only to ± 0.1 magnitudes. Pascu concluded by saying that spacecraft pro-

vided positional results comparable to the best Earth based astrometric observations. The sparsity of the spacecraft data is ample justification for the continuance of telescopic observation.

Orbits

G. A. Wilkins (Royal Greenwich Observatory) discussed the orbits of Phobos and Deimos making special reference to the way in which these are varying with time. Deimos orbits Mars in 30 h 18 min and Phobos in 7 h 39 min, Phobos having the strange distinction of being the only satellite in the Solar System which orbits its central body faster than that body spins on its axis. Looking at the satellites from the surface of Mars Deimos would rise in the East and set in the West just like our Moon does but Phobos rises in the West and sets in the East. There are two other strange things about the orbit of Phobos. First its mean distance from the centre of Mars is 9,400 km whereas the Roche limit is about 8,700 km (any fluid body getting closer to Mars than 8,700 km will be pulled apart by differential tidal forces). Second, it is getting closer. This automatically increases its mean orbital velocity V , Sharpless (*Astr. J.* **51**, 185; 1942) concluding that $\Delta V/V$ per period was about 8×10^{-12} . This secular acceleration could be caused by a resisting medium in the region of the satellite orbit (the outer remnants of a Martian atmosphere?) or be a consequence of tidal friction, the absence of seas on Mars however requiring this to be a body tide. The Soviet astrophysicist Shklovskii in the 1950s favoured the former explanation postulating that the large $\Delta V/V$ was due to the low density of Phobos; he suggested a density of $10^{-3} \text{ g cm}^{-3}$. Shklovskii astounded the scientific world by suggesting that Phobos was in fact a hollow, artificial satellite produced by an advanced Martian civilisation.

Wilkins was not convinced that present-day analysis of observational data indicated that there was any secular acceleration taking place. Spacecraft data seemed to indicate that it was at least a factor of three smaller than that deduced by Sharpless. Even at this lower rate Phobos would crash on to Mars in about 10^8 years assuming of course that it was not pulled apart into a myriad of rocky pieces on coming closer than the Roche limit. Such a disruption would give Mars a ring system like Saturn and Uranus.

Spacecraft observations of the surface features of the Martian satellites were discussed by Jo Veverka (Cornell University) (see also *Scient. Amer.*, February, 1977). In 1971 Mariner 9 obtained 29 views of Phobos and 9 of Deimos with an average resolution of 200 m. Surprisingly both satellites were

found to have almost identical shapes, that of triaxial ellipsoids, Phobos having principal diameters of 27, 21 and 19 km, Deimos of 15, 12 and 11 km. Why this is so is still somewhat of a mystery. However, Don Gault (Ames Research Centre) found that high velocity impacts on spherical targets produced a crater at the impact point, spalling on the opposite side of the sphere and often the complete removal of the entire outer layer of the target. The inner core left behind had a shape surprisingly like that of Phobos and Deimos. Both moons rotate synchronously with their orbital periods and their long axes pointing directly at Mars.

View from Viking

The moons are not the same colour as Mars. Mars is red, they are blackish-grey, indicating that the surfaces have different compositions. Albedo and spectral reflectance measurements indicate that the satellite surfaces are made of C1 carbonaceous chondritic material similar to the Orgueil meteorite and the asteroid Ceres. This material has a density of about 2.0 g cm^{-3} . Crater counts indicate that the surfaces are at least 2×10^9 yr old and Veverka concluded that the satellites must have been formed in the asteroid belt and have moved to Mars later on. The Viking spacecraft has approached as close as 100 km to the satellite surface and has obtained pictures with a resolution of about 10 m. These pictures reveal curious rill-like fracture patterns around Stickney, a 10 km diameter crater on Phobos. (This crater has a diameter about 40% the maximum diameter of Phobos itself!) These could be failure cracks produced in the rocky satellite when the crater was formed. Crater chains have also been found, these being parallel to the orbital plain (which is also the equatorial plain of the satellite). These could be formed by secondaries blown off the satellite by impact events, these secondaries having velocities lower than the escape velocity from the Martian gravitational field. After orbiting the planet for a time (that can be as long as 50 yr) some of these fall back on to the satellites, forming secondary craters. The surface also contains peculiar groove-like striations, these forming segments of a series of small circles which are all perpendicular to the Mars-satellite line. Tidal mechanisms are being considered as a possible source of these.

W. Michael (Langley Research Centre) discussed the changes induced in the orbit of the Viking space probe by its close approach to Phobos. These indicate that Phobos has a mass of $1.1 \times 10^{19} \text{ g}$ (about 2×10^{-8} the mass of Mars). The uncertainty in this value is

about $\pm 15\%$ and coupled with a volume which is also known to $\pm 15\%$ this gives a density with a probable value between 2.1 and 2.2 g cm^{-3} and an error of about $\pm 20\%$.

The problem of the origin of the Martian satellites was discussed by M. Woolfson (York University). Theories are legion—a usual case when there is not enough available evidence. Most have the two satellites originating at the same time mainly because their orbital orientations are similar, as are their photometric characteristics. They are thought of as being planetismals left over from planetary formation, condensed filaments produced by collapsing protoplanets, captured asteroids that crawled into the inner Lagrangian points then lost energy by tidal friction giving them more stable orbits, fragments from a disrupted asteroid that passed within Mars' Roche limit, fragments from a collision between two asteroids that took place within Mars' gravitational sphere of influence or fragments produced by collisions between protoplanets or between protoplanets and Jupiter, these collisions resulting in all the asteroids and terrestrial planets and their satellites.

So a hundred years has seen a vast increase in our knowledge of these two fascinating satellites. One of the bonuses of sending a spacecraft to Mars is that you have these asteroid-like bodies on hand for investigation if Mars becomes blotted out by dust storms. Viking went within 100 km, let us hope that in the near future a spacecraft will be sent to land on Phobos or Deimos to see what their surfaces are really like. □



A hundred years ago

THE annual meeting of the Institution of Mechanical Engineers opened on Tuesday at Bristol. Mr. T. Hawksley, C.E., in his opening address, said it was the duty of the government to adopt such timely measures as would secure to us the paths of the ocean for our food inwards and our manufactures outwards. He deprecated the building of enormous and unwieldy floating castles, and advocated the construction of a fleet of swift, light, well-engined ships, equally capable of sailing or steaming. He thought the extreme action of some of the working classes the cause of England's trades going abroad.

From *Nature* **16**, 26 July, 256; 1877.



Prospects in oceanography

At a time when science in general has lost its magic, the science of the earth—and especially of its mysterious ocean—has been increasingly appreciated for its combination of difficulty and importance.

Plate tectonics helped a lot by providing a unifying hypothesis which could be readily grasped, and which justified the investment in earlier observations at sea, especially of the magnetic field. The deep drilling programme, now IPOD, has been a triumph of science, technology and organisation. The physical oceanographers have learned how to make direct measurements of ocean currents and discovered that the deep ocean is quite different from what they expected. There are eddies which may well be the marine equivalent of atmospheric depressions, longer lasting but smaller and slower. If they resemble atmospheric depressions to the extent of dominating the poleward transfer of heat they will be vital to our understanding of climate and climatic change. Marine biologists can now unambiguously sample macro-plankton in the deep ocean; zoogeography is becoming a quantitative science at last. Marine chemistry is now a subject in its own right.

Much of this progress has been made possible by instrumental development and skilled engineering. We have exchanged our sextants for satellites, we can dive into rift valleys, we can image the earth from space and make micro-photographs at fantastic magnification. And yet we have only just started to understand the ocean; it remains dark, mysterious, enigmatic, the last great pit of ignorance on our planet. There are simple-sounding questions to which we cannot yet give satisfactory answers. An engineer designing an offshore structure may feel that we should know how surface waves are generated by the wind. A forward-looking fisherman might expect us to know about the life-cycle of krill. A climatologist might think that we should know how much heat the Gulf Stream transports.

Oceanographers used to form a small group of rather eccentric, half-scientist, half-seaman explorers. Everybody knew that the ocean provided a support for ships, a source a food, a sink for waste and a regulator of climate, but it wasn't thought that science could help the engineers much. In those days, not so long ago, anyone making a long series of wave-observations would have been accused of wasting taxpayers' money. Now he is asked why he hadn't made many more.

There is an inevitable mismatch between the timescale of marine research and of the practical problems to which it can contribute. Environmental problems arise unexpectedly: stamping them 'immediate' cannot reduce the time it takes to make the necessary observations or get some fundamental understanding. Oceanographers must, therefore, gather as much basic information as they can: they cannot predict, any more than anybody else, what problems will come up so they must necessarily work on a variety of problems, trying to connect them into a self-supporting generality from which special cases can be sensibly studied.

The money spent on such basic research is perhaps analogous to a life assurance premium to guard against an inevitable contingency rather than to react to a particular, predictable event. It is generally desirable to pay one's assurance premium regularly and over a long period. Marine research is a bit too long-term to suit most of its customers, to use the present jargon. Perhaps it needs patrons rather than customers; in the long run, as Lord Rothschild probably appreciates, it is selective patrons who get rich.

A few problems are, however, so long-term as to be foreseeable. One concerns the disposal of high-level radioactive waste. It might be placed on, or preferably in, the sea floor; a study of the effects is clearly essential. Fortunately much relevant information and expertise is already available, thanks to an investment in basic research, so progress can be made fairly quickly, which in this context means in a decade or two.

Another long-term problem is climatic change. Man is vulnerable to relatively small changes in his environment; there is growing concern that we know so little about the mechanism of climatic change, that we cannot assess whether, or to what extent, man's activity contributes to it. If anything is clear it is that the ocean is an important component in the system—the pioneers of air-sea interaction and of ocean circulation studies weren't frittering away the taxpayers' money after all. There is a lot left to do—and it will take a long time.

Long-term, the outlook for oceanography seems good. The short-term gloom stems from a basic incompatibility between an indivisible ocean and an increasingly nationalistic world, which is offsetting what some find the main attraction, the combination of personal effort with international cooperation. Sensible legislation is clearly desirable and able scientists and lawyers are working on it. But it does not seem likely that United Nations Conferences on the Law of the Sea can do much to create the goodwill needed, whatever is finally agreed. The badly needed bridge between politics and science can only be filled by the slow process of education. It may well be that training and mutual assistance in marine science is the most valuable contribution some mature oceanographers can now make. There is a growing risk that marine research will be slowed down, not so much by lack of access but because the bright young people on whom progress depends will prefer a branch of science where freedom to make observations does not depend on unpredictable and unexplained bureaucratic decisions.

Some may argue that a check to the progress of marine research would be a justified casualty of peace, a transient on the way to a new world order. It seems much more likely that oceanographers will be in the lead.

Henry Charnock

Professor Charnock is Director of the Institute of Oceanographic Sciences, Wormley, Surrey, UK.

Developments in the Law of the Sea

Edward Miles

Institute for Marine Studies HA-35, University of Washington Seattle, Washington 98105

In this and the following article developments in the law of the sea are described and their impact on marine research is considered. Agreement seems close on much of the negotiating text concerned with coastal and economic zones. But beyond, on the sea-bed in deep water, there are still many difficulties to be resolved. Is the outlook for scientific research therefore bleak? It is argued that if scientists come to terms with the new order, prospects for research in the economic zone and beyond could well be bright.

United Nations Conferences on the Law of the Sea and Associated Meetings

First UN Conference on the Law of the Sea (UNCLOS), Geneva, 25 February–27 April 1958.

Second UN Conference on the Law of the Sea, Geneva, 17 March–26 April 1960.

Ad hoc Committee to Study the Peaceful Uses of the Sea-bed and the Ocean Floor beyond the Limits of National Jurisdiction. (Established by UN General Assembly resolution 2340 (XXII) of 18 December 1967.) Three sessions, New York and Rio de Janeiro, March–August 1968.

Committee on the Peaceful Uses of the Sea-bed and the Ocean Floor beyond the Limits of National Jurisdiction. (Established by UN General Assembly resolution 2467 (XXIII) of 21 December 1968.) Eleven sessions (plus additional meetings), New York and Geneva, February 1969–August 1973.

Third UN Conference on the Law of the Sea. (Established by UN General Assembly resolutions 3029A (XXVII) of 18 December 1972 and 3067 (XXVIII) of 16 November 1973.) First session (administrative matters only), New York, 3–15 December 1973; Second session, Caracas, 20 June–29 August 1974; Third session, Geneva, 17 March–10 May 1975; Fourth session, New York, 15 March–7 May 1976; Fifth session, New York, 2 August–17 September 1976; Sixth session (scheduled), New York, 23 May–8 (or 15) July 1977.

In the latter half of the twentieth century the strategic significance of the ocean has substantially increased as a result of the relative invulnerability of the submarine-based deterrent. At the same time, the incidence of violent conflict over purely ocean-related issues has been quite low. But, since the first attempts to codify the Law of the Sea under the auspices of the League of Nations in 1930, there has been a dramatic increase in non-violent conflict over ocean resources, both living and non-living.

The most contentious issue of the period 1920–60 was the limit of the territorial sea as it pertained to jurisdiction over coastal fisheries. This, in turn, was linked to the issue

of freedom of passage through straits used for international navigation since such passage was seen to be threatened by the extension of coastal state jurisdiction over fisheries. In 1968–69, the major problem in the UN Sea-bed Committee was the question of limits to the continental shelf and jurisdiction over the ocean floor beyond the shelf. In 1973–77 at the Third United Nations Conference on the Law of the Sea (UNCLOS III), the shelf and sea-bed remain contentious issues but the number of controversial problems has proliferated. These include the exclusive economic zone concept, straits used for international navigation, archipelagos, islands, scientific research, ship-generated pollution and a host of other things.

The primary function of the law of the sea has always been to arrive at some balance between the exclusive interests of a few states relative to any particular use of the ocean in various spatial locations and the inclusive interests of the rest of the world system. Since 1930 the search for a globally acceptable balance has been posed increasingly in terms of choosing between the inclusive interests of free navigation and the exclusive interests of extending coastal state control over living and non-living resources in wider areas of the water column and continental margin adjacent to their coasts. It is estimated that the focal point of the current negotiations, the 200-mile exclusive economic zone, would result in an additional 35%–37% of the landmass of the planet passing under national control.

The conferences of 1958 and 1960 differed from that of 1930 in the number and type of participants. Twice as many countries (86 and 88) participated in the latter as in the former (44). Expectations about the possible contributions of marine technology and ocean resources to national development seem to have been significantly stimulated at this time. In 1974 participation in UNCLOS III became almost universal with 137 participating countries, out of the 149 invited. In 1976, the number of participants increased to 156. The alignments in the UN Sea-bed Committee from 1968–73 and in UNCLOS III itself were no longer those of the General Assembly but resembled the alignments of the United Nations Conference on Trade and Development (UNCTAD) more and more. In doing so they foreshadowed Assembly politics of the 1970s with the emphasis on the confrontation between advanced industrial and developing countries. This was clearly the most complicated, and perhaps the most difficult, law-making conference the United Nations had ever hosted. There were many more countries represented than ever before; consequently many more interests had to be reconciled. Many of these interests seemed mutually exclusive. Given the primacy of the North–South confrontation, no single group seemed capable of providing effective global leadership.

Technical experts, though in attendance, were tightly controlled by their delegations. They were not allowed to have any independent role as in 1958, since science itself had come under heavy attack. There was no longer any trust and repeated assertions about the necessity of facilitating disinterested scientific investigations of the world ocean

were consistently denied. This was a fight to redistribute ownership and control over world ocean resources and science was seen to be a weapon in this fight.

States differ both in the variety of ocean-related capabilities they command and in the marine biogeophysical conditions which they face. These differences are of particular importance in explaining much of the political behaviour and outcomes in the negotiations. But, in addition, the process and the results of negotiating are very significantly affected by the dynamics of conference diplomacy in global settings. These dynamics include the need to build and maintain 'winning' coalitions as defined by the majorities in the Rules of Procedure, the method of arriving at decisions, the extent to which purely ocean-related issues are affected by considerations external to the oceans, and many other factors.

Emergence of a new regime for the oceans: content and process

Content: The heart of the new regime is contained in two documents: the Revised Single Negotiating Text (RSNT), Parts II and III. Although a few issues remain to be resolved, the outlines of the new law of the sea are in fact quite clear. The situation relating to the sea-bed beyond national jurisdiction, however, presents grave negotiating difficulties and threatens the whole conference with official failure, that is, no treaty. Consequently, the RSNT, Part I, is at the moment extremely controversial.

The basic ingredients of the package in the RSNT, II and III, are as follows. (1) A 12-mile territorial sea with separate regimes governing innocent passage and passage through straits used for international navigation. (2) An exclusive economic zone of 188 miles beyond the outer limit of the territorial sea wherein the coastal state exercises sovereignty over living and non-living resources, exclusive jurisdiction with respect to other activities like scientific research, and jurisdiction with respect to yet other activities like preservation of the marine environment. (3) The coastal state exercises sovereign rights over its continental shelf, which is defined as either the continental margin (including the shelf, the slope and the rise) or 200 miles, whichever is greater. (4) Archipelagic states may draw straight baselines connecting the outermost points of the outermost islands and drying reefs, before delineating the outer limits of the territorial sea, economic zone, and continental shelf. (5) Island states may extend jurisdiction as described but "Rocks which cannot sustain human habitation or economic life of their own shall have no exclusive economic zone or continental shelf" (Article 128, paragraph 3).

The RSNT, III, is itself divided into three parts. The first part deals with preservation of the marine environment and contains primarily very general obligations for states relating to pollution from land-based sources, from or through the atmosphere, and by dumping. The most controversial section (VII) is contained in Articles 23-42 dealing with the enforcement of these provisions over ships and particularly over vessels operating within the economic zones of coastal states. The source of controversy is that the issue is linked to the status of the economic zone as it affects navigation and therefore the security interests of particular maritime states. In a similar fashion, the part on marine scientific research also has broader implications but the specificity of coastal state control is greater. There are, however, some incompatibilities between the approaches adopted in Parts II and III.

With respect to the sea-bed issue, the major points of controversy are as follows. (1) The basic principle governing

exploitation of the area and its resources. The Group of 77 insists that there be only two modalities: either the International Authority directly and exclusively exploits the resources and/or the Authority exploits in conjunction with other contracting parties via the mechanism of joint ventures. On the other hand, the advanced industrial countries insist that a third modality be allowed, that is, exploitation by contracting parties alone. (2) The detailed rules and regulations governing the relationships between the Authority and contracting parties. (3) The structure, functions and powers of the International Sea-bed Authority and, in particular, the method of making decisions within the Council. (4) The method of financing the Enterprise which is to be the operating arm of the Authority. And (5), the question whether the Authority is to be given the competence to set its own resource policy; this implies the capacity to impose production and price controls on contracting parties.

Process: From the first substantive session in 1974, the Conference has suffered from the problems of size and complexity incurred in attempting to negotiate an agenda of more than 100 issues simultaneously among 137-156 nation-states. These difficulties have been aggravated by the Rules of Procedure which emphasise the necessity of achieving consensus, the dynamics of global conference diplomacy which are often at odds with the distribution of interests in the substantive problems being negotiated, and the considerable natural inertia which such a social system brings in its train.

Even during the preparatory phase (1971-73), it was clear that there would be two major overlapping dimensions to the conflicts which would arise in the Conference. The first, already mentioned, would pit developing countries against advanced industrial countries. The second, however, would cut across the former by pitting all landlocked and geographically-disadvantaged countries against those coastal states which stood to gain most from the exclusive economic zone concept. The combined effect of these two factors was to increase the probability that no coalition would actually have the votes to achieve a two-thirds majority and therefore significantly to delay the emergence of a treaty.

The second major problem of the negotiation process at UNCLOS III concerns the sea-bed issue in Committee I. This was in reality a four-dimensional problem. The first dimension was that only 20-30 countries actually had substantive interests at stake in the negotiations. For the remaining 100 or so delegations the symbolic issues were paramount and this was the issue on which the North-South confrontation was at its most extreme. The second dimension is represented by a group of states, led by Algeria, seeking to make this issue into a central issue of the New International Economic Order. They were ultimately successful in this strategy and were significantly aided in doing so by the third dimension: a very serious internal split within the US delegation between the Departments of the Treasury and Interior over the line that would be followed. The fourth dimension was an idiosyncratic one but very significant, nonetheless, since it involved a severe personality conflict between the chairman of Committee I and the chairman of Committee I's Working Group. All of these dimensions affected the negotiations in important ways which we do not have the space to describe in detail.

One final point should be made about the process of negotiations in Committee I. As a result of the scope and technical complexity of the issues being dealt with, each major issue was negotiated separately rather than in a total package as in Committee II with the economic zone and transit through straits. Consequently, the same confrontations occurred at every stage and increased the difficulties of reaching agreement.

Probable outcomes and their effects

Will the Conference be able to produce a treaty or not? This depends primarily on whether either the advanced industrial countries or the Group of 77 capitulate on the major issues of access to and control of the international sea-bed area. If the United States capitulates, it is not clear at the moment that either the EEC countries, Japan, or the US Senate will concur. If there is continued stalemate, the Conference will be by-passed by unilateral action of the US Congress probably in late 1977 or early 1978. In that event, undoubtedly the EEC countries and Japan will follow suit.

While the sea-bed is the primary problem for the next session (May–July 1977), there are still three difficult issues remaining in Committees II and III. (1) The status of the economic zone or whether the zone is to be declared waters under national jurisdiction with specified easements for coastal states or vice-versa. If neither is the case, the problem is whether the zone is unique. The underlying problem concerns the residual rights of navigation within the zone and their significance for the defence establishments of the USA and USSR. (2) The problem of coastal state control of scientific research undertaken within its economic zone. (3) The continued, intense opposition of the Group of landlocked and 'geographically-disadvantaged' States to the articles establishing the economic zone. While

these are difficult problems, there are indications that solutions can be found.

If the Conference breaks down over the sea-bed issue, the result is likely to be continued controversy over the RSNT, I, but a high degree of acceptance of the provisions contained in the RSNT, II and III. Re-revised versions of these are likely to be presented as formalised President's Texts outlining the new regime for the oceans. Already, in the large number of unilateral extensions of coastal state jurisdiction of 200 miles, most of them are obviously patterned after the RSNT, II.

What then are the likely substantive effects of such a regime? The most sweeping changes would relate to fisheries wherein almost all the world's demersal species currently fished and about 60% of tuna and tuna-like species will pass under national control. The effects on management of tuna in the short term will be grave since control over the stocks will be fragmented. The stability of the regime for navigation will be in question as far as the issue of coastal state control over ship-generated pollution in the economic zone is concerned. The scientific research activities of advanced maritime countries, and particularly of the USA and USSR, in the economic zones of other coastal states are likely to be adversely affected. Finally, it will become increasingly more clear that the economic zone will be a multi-purpose zone for which a capacity to provide co-ordinated management across all activities should be developed.

Implications of the Third United Nations Conference on the Law of the Sea for marine scientific research

Desmond P. D. Scott

Intergovernmental Oceanographic Commission, Unesco, 7 Place de Fontenoy, 75700 Paris, France

THE Third United Nations Conference on the Law of the Sea (UNCLOS) and its predecessor—the United Nations Committee on the Peaceful Uses of the Sea-bed and the Ocean Floor beyond the Limits of National Jurisdiction—have, for the last 10 years or so, from a starting point of Arvid Pardo's 'Common Heritage of all Mankind', been steadily trying to reach agreement as to how the resources of the oceans beyond the limits of the old territorial sea—a cannon shot range of 3 miles from the coast—can be divided equitably between the 150 odd sovereign member states of the United Nations.

Numerous other problems have arisen along the way which have required investigation and solution but the resources, both living and non-living (mainly the latter), are the sole real reason for the whole exercise.

This monumental task which is as yet far from complete, seems to be moving at a desperately slow pace but it should be remembered that it is indisputably the greatest single attempt ever made to develop an entirely new international legal regime for a highly complex subject area having numerous facets and linked interests, all of which have to be taken into account. When complete the resulting Convention will cover the waters of the oceans, the sea-bed and the subsoil beneath, over an area covering more than 70% of the surface of the globe. It should be realised that the First UN Conference on the Law of the Sea (the Second

was limited in scope and was a failure) was, in essence, an attempt to codify existing traditional laws that had developed haphazard over the years, but that this is an entirely different type of exercise that cannot be carried out in a hurry if the final result is going to be acceptable to a majority of the countries of the world.

It has been said that the Conference is an attempt to avoid the development of a situation similar to the African land grab of the nineteenth century but in the event, this is exactly what has occurred but with more contenders for the prize and no indigenous peoples to cause difficulties in later years.

The preceding article describes the way in which the Law of the Sea has developed, so no attempt has been made herein to identify the various draft articles and discussions dealing with marine scientific research. But, there are a number of basic agreements that have now been reached and are certain to become generally accepted law in the next few years whether or not an acceptable Convention is forthcoming from the Conference; the likely effect of these on the future of the marine sciences is considered below.

Crisis in marine research

It cannot be denied that fundamental marine scientific research and much applied research of vital importance to developing as well as developed countries, is now facing a

crisis quite unlike that ever faced by any other science. Ocean processes and living resources do not respect boundaries arbitrarily imposed by man, and scientists are finding that their floating laboratories can no longer be sited where they are needed, nor do they have unrestricted access to areas where they can collect the data that are essential for their research work.

This is inevitable during the transitional period, that is, the duration of the Conference, and it will doubtless cause a temporary withdrawal to accessible areas—to waters under the jurisdiction of like-minded coastal states and to the deep oceans, though the legal status to be accorded the deep ocean floor and subsoil is still under discussion.

There is every indication, however, that coastal states have been stimulated by the Conference on the Law of the Sea and the number of enquiries and requests from developing coastal states for help in building up their marine science infrastructures is growing to flood proportions. Countries are turning in increasing numbers for direct help by bilateral aid and to such bodies as Unesco, OECD and the EEC.

In the long term therefore, it is already abundantly clear that there will be no insurmountable restrictions on the marine sciences. The length of the present hiccup will depend entirely on the perception of scientists in the limited number of countries with developed and sophisticated marine science infrastructures, and their willingness to help other coastal states to develop their expertise and capabilities as part of their on-going marine science programmes, and not as a separate exercise which is always the responsibility of another government department.

Economic zone and beyond

The great majority of participants at the Conference are now agreed (as noted in the preceding article) that the territorial sea should be increased in width to 12 miles and that there should be an economic zone of 200 miles width (from the coast) over which coastal states would virtually have complete sovereignty so far as living and non-living resources are concerned. Furthermore, where the geographical continental shelf extends more than 200 miles offshore, the coastal states will still have the same sovereignty, though a proportion of the profits from exploitation of resources in this extended area may have to be used for the benefit of developing countries. The single most important matter of concern to marine scientists is the almost unanimous wish of the developing countries to insist on scientific research and exploration in these areas being placed in the same category as exploitation of resources, an activity that can only be carried out with the consent of the coastal state. This attitude is due to the not unfounded belief that knowledge of the existence of resources leads to a position of power and a very well founded fear that research ships do not always confine their activities to the project declared and may be collecting data detrimental to the coastal state, coupled with the fact that it is well known that military and commercial interests are in most countries given precedence over those representing scientific research.

The 200-mile width chosen for the economic zone is an arbitrary figure based originally on claims made (and still maintained) by a number of South American countries for the width of their territorial sea. This had nothing to do with the geographical concept of the continental shelf (which had been the basis of the rights of national jurisdiction contained in the 1958 convention, on the principle that the continental shelf is an extension of the continental land mass) but rather was based on the economic need of developing countries on the west coast of South America to lay claim to what was then developing into the greatest fishery resource in the world.

There is no question, however, but that the deployment of scientific research ships in the strip of ocean 200 miles in width from the coast and beyond that distance to an as yet undefined limit which will certainly be in the deep ocean, will be firmly under the jurisdiction of the coastal states. The most likely contender for this outer limit seems to be the outer edge of the continental rise but this is so ill-defined in most oceans that it will remain a theoretical limit rather than one that can be plotted on a chart, for many years to come.

Where the limits can be drawn in without too much difficulty, and future jurisdiction is reasonably clear, is in the regional seas, in particular the North Sea, Baltic, Mediterranean, Caribbean, South China Sea, etc., where the whole sea will virtually be under the jurisdiction of the coastal states.

In the Mediterranean, for instance, this situation has already been recognised quite clearly by the coastal states of the region, and they have retained for themselves the right to formulate all policy regarding research on the marine environment being developed under the auspices of the United Nations Environment Programme (UNEP). This does not, however, prevent states from outside the region, in particular developed countries with research ships and appropriate expertise, participating in the implementation of such research projects.

Prospect for more cooperation

At first glance this might seem to mean that the only research work that will be permitted in such a region and off the coasts in the open ocean, will be that having immediate application to the problems and interests of the coastal states and, as many of these states will be developing countries with little interest in sophisticated research, the outlook may superficially look bleak for fundamental and other scientific research work having wider implications than a purely local interest and application. But this ignores the basic bargaining power in the hands of the developed countries of trained manpower and ships which, if used properly, will allow both parties to obtain what they need: the developing countries—training and education in the marine sciences, experience and strengthening of their infrastructures, and the developed countries—the right to work in areas of scientific interest. Such an arrangement will only work if each party recognises the interests of the other and is prepared to cooperate closely to their mutual benefit.

This calls for three quite separate types of international education to meet the changing situation.

- (1) Education of scientific administrators and decision makers in developing coastal states to enable them to understand the peculiarities inherent in the administration of marine science.
- (2) Education of senior scientists and scientific funding agencies in developed countries to get them to face up to the changing situation and abandon their rearguard action for freedom of scientific research and separate definitions (and therefore regimes) for fundamental and applied research.
- (3) Education and training of young marine scientists in developing countries to build up the national infrastructures and know-how in these countries, so as to educate out the suspicion now so rife in the UN Conference on the Law of the Sea and in national dealing between developed and developing coastal states.

Let us take these three basic requirements and look at them more closely. (1) The second half of the twentieth century is the era of 'development'. The United Nations Development Programme has been established and many developed countries have set up development agencies to provide aid to developing countries. Such agencies usually

work through bilateral agreements, sometimes using the UN specialised agencies in the process. At the same time, the developing countries are clamouring for what is known as 'transfer of technology', as if technology consists of something that can be handed over in the form of a package bringing immediate affluence to the recipient. In fact, much of the money poured into 'development' over the years has been wasted due to two basic factors: first, lack of basic research before the actual development project is started; and second, inadequate training of personnel in the developing country to enable the project to be undertaken in the first place using local skilled personnel and, even more important, to ensure that the activity whatever it may be, continues in full operation when the basic project is complete and external support, both in the way of funds and personnel, is withdrawn.

This is just as applicable in the field of marine science, and marine affairs generally, in which the real lack of most developing countries is not technology as such and frequently (particularly in recent years) not even funds, but trained personnel and a viable infrastructure bringing together all the various interests—universities, navy, fisheries, basic research interests, marine engineering and so on.

Once an educational system has been devised that will produce a regular supply of trained personnel on the market, a country is able to absorb outside technological know-how and adapt it to its national needs. In the field of marine science, the best example of a developing country whose administrators and decision makers have realised this truth, is Mexico which is well on the way to becoming self-sufficient in marine scientific manpower and therefore a very powerful nation in marine affairs generally.

There is a real need to educate scientific administrators and decision makers in developing countries to the fact that there is no short cut to the wealth of the ocean's resources and that their highest priority should be the development of a cadre of trained manpower.

(2) There is no doubt that coastal-state control over scientific research in the economic zone is here to stay but there are far too many senior scientists unable to assimilate the fundamental political changes that have taken place in their chosen area of activity. Fighting the inevitable can do more harm than good as it can only strengthen the belief, sincerely held in many countries with no fundamental scientific background, that scientific research in the oceans is a cover up for other activities which would be to their detriment.

In the past, senior scientists planning research cruises have as a concession offered a few places aboard ship during a cruise to one or two scientists from the coastal states off which they wish to operate. Developing country scientists that have participated in such cruises have almost unanimously declared that they have been valueless both to themselves and to their country, as they have come unprepared to a sophisticated activity run by a research group having close personal relations and a full knowledge of the work being conducted; as a result, they have found themselves as outsiders with little understanding of what is going on.

In the future it must be recognised that the research is being conducted in what is virtually the national territory of the coastal state, just as if it were taking place on that part of the country which is above water. As a result the coastal state, through its marine scientists, now has the full right to participation in the research project from start to finish. This means that the coastal state scientists must be fully involved in the project from the first planning stages, through the data collection/cruise phase, to the data analysis phase and the assessment and writing up of the results. Should it be more convenient,

due to availability of facilities, development of equipment, and so on, to carry out the first and last phases in a laboratory in a developed country, for example of, that providing the research vessel, this is where the developing country scientists will have to work at full cost to the developed country. In fact the trend should be towards developed country scientists setting up facilities in developing country laboratories (at developed country expense), in order to help developing countries to strengthen their own facilities.

Whilst the developed country scientists have got to recognise the essential change in their working conditions if their science is not to die, the scientific funding bodies and aid agencies must come to realise that they are in this together and that this is a joint activity within the mandate of both bodies. The first reaction of a well known funding agency to such a suggestion was "we cannot afford it" and no doubt initially aid agencies would have a similar reaction against what they would consider to be funding national research projects, even if in the waters of countries they are set up to aid. Such attitudes must be modified drastically in the recognition that this is not only to the benefit of both in the long term but is basically the only solution if global marine science is to survive. A logical development would be to bring the two activities under the same administration and no doubt this will be the final outcome in many countries, but such moves will clearly only come about in slow time.

(3) Finally, there is the widely recognised need for basic education and training of young marine scientists in the developing countries (as touched upon in (1) above).

There are many difficulties inherent in the development of a viable infrastructure in a country with little or no tradition in the marine sciences, such as wastage and 'brain drain', which would be out of place if discussed in this article, but building up a strong cadre of marine scientists in each coastal state is the only way of overcoming the almost pathological aura of fear and suspicion that has grown up in many developing countries with regard to marine scientific research. It will only be when countries have their own trained scientists to whom the politicians can turn for advice, or, even more important, who can apply pressure on the politicians to take action, that this fear and suspicion will subside. The intergovernmental bodies will have a growing role to play in this regard if they can be so structured to gain the trust of the developing countries but they can never replace a strong national cadre of trained marine scientists.

It will be noted that education and training of marine scientists in developed countries has been omitted from this list. This is because it is considered to be a national problem for the developed countries concerned. It must however be clearly recognised, and soon, that education and training of marine scientists in developing countries is an international problem requiring urgent and greatly increased action.

Bright future?

I am convinced that the future for the marine sciences is bright but that this will depend on two basic factors: first, the end of the UN Conference on the Law of the Sea and a stable period during which the results and implications thereof can be assessed and put into practice; and second, the recognition and acceptance of a new and very different international regime for marine science in the future, by those who have been brought up to believe in their historical right to conduct marine scientific research where and when they please (subject only to finance). When these two factors have been met, such a regime could lead to accelerated progress in the marine sciences through a greatly enlarged global community of scientific colleagues working towards common scientific goals.

1967–77: Ten years of marine geology and geophysics

Xavier Le Pichon

CNEXO, B.P. 107, Paris 75116 and COB, B.P. 337, 29273 Brest, France

Plate tectonics gave a great stimulus to marine geology and geophysics; in the years since its enunciation there have been major drilling programmes, a new science of global palaeoenvironmental oceanography and much more detailed oceanographic surveys. The pace shows no sign of slackening—in the next few years the passive continental margins will be much studied.

In April 1967, Jason Morgan¹, having unexpectedly changed the subject of his talk, demonstrated to a large audience at the American Geophysical Union meeting, in Washington, how sea-floor spreading in the Atlantic Ocean had occurred between two rigid 'crustal blocks' moving apart on the surface of the spherical earth. This 'plate tectonics' model of the Earth's surface was implicit in the sea-floor spreading hypothesis of Hess², with its two corollaries of Vine and Matthews³ magnetic lineations and Tuzo Wilson⁴ transform faulting. But although the magnetic anomaly profiles across the East Pacific rise had convinced many earth scientists by early 1966 of the reality of sea-floor spreading, the demonstration made by Morgan was largely ignored by the audience. It is only during the following years that the full impact of rigid plate kinematics led to the present revolution in the Earth sciences. And this revolution has been most spectacular in marine geology and geophysics, which, with earthquake seismology, furnished most of the discoveries on which it is based.

It is in the field of global kinematics, past and present, that the progress has been most rapid. Because plates are rigid, it has been possible to construct a fairly precise global quantitative model for the present interaction of the 11–12 major plates, along their boundaries which are generally covered by several kilometers of water. In addition, it has been possible to decipher, at least in a preliminary way, the complex kinematic history of all large oceans, through the powerful Vine and Matthews method of dating the igneous oceanic crust. Thus, for the first time, there is a successful quantitative model of the dynamic evolution of the earth surface, a model which has been repeatedly tested, in particular through the use of deep-ocean drilling.

Deep-sea drilling

In 1968, at the time when the plate tectonics model was being accepted by the scientific community as the working model, the D. V. Glomar Challenger set out to drill her first hole through the floor of the Gulf of Mexico. She is now drilling her 418th hole and the results⁵ obtained by the Deep Sea Drilling Program (DSDP), under the sponsorship of JOIDES (Joint Oceanographic Institutions for Deep Earth Sampling) have shown that the proposed kinematic evolution of the oceans is supported by the analysis of the sedimentary column and the underlying igneous crust. In particular, the magnetic reversals absolute timescale, which is the basic tool to unravel the history of the oceans and which is now known from present to Middle Jurassic (165 Myr ago) with a relatively good precision (≈ 10 Myr), was entirely calibrated by deep-sea drilling results.

The single most important accomplishment of deep sea drilling, however, was the creation in the last 3–4 yr of a new field of science—global palaeoenvironmental oceanography. This occurred as it was progressively realised that the drastic changes in the Earth surface configuration are highly correlated to changes affecting the hydrosphere and the atmosphere, but also the biosphere. It now seems possible to obtain reasonable physical models which broadly account for the deterioration of climate during the last 40 Myr, simply by taking into account the different ocean–continent distribution pattern (for example, ref. 6).

Palaeoenvironmental oceanography

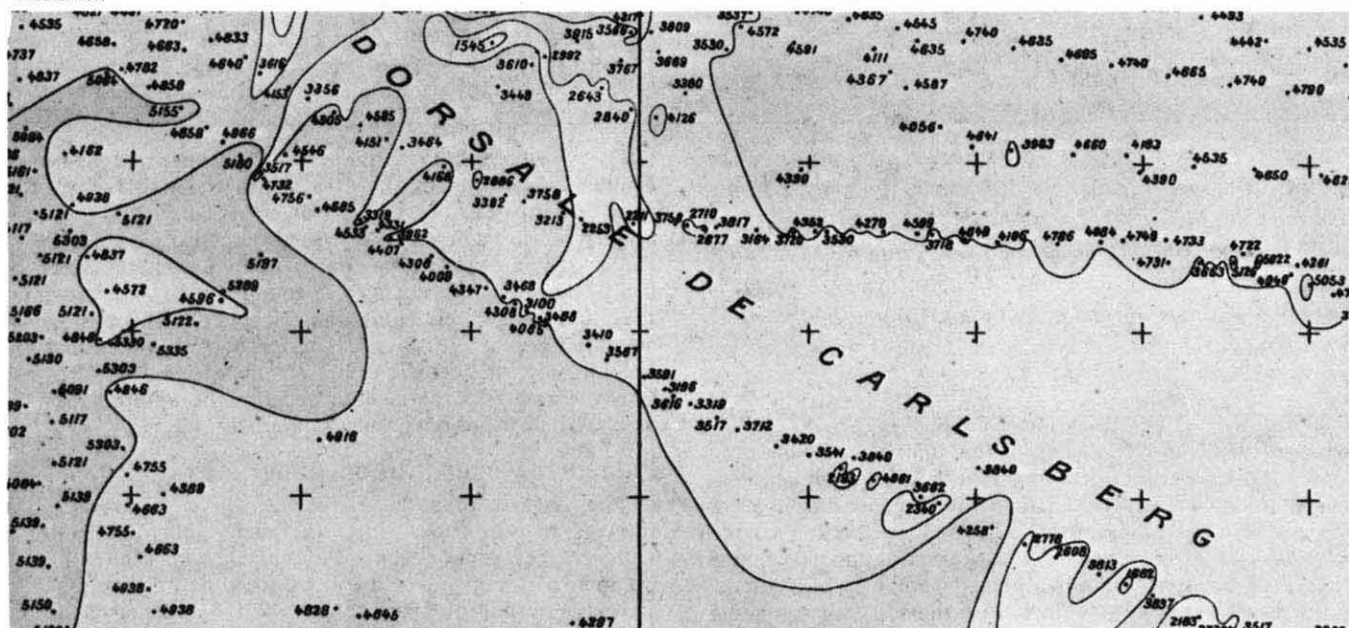
In the last 200 Myr the Earth surface configuration has changed from a supercontinent, Pangea, surrounded by a proto-Pacific super-ocean, Panthalassa, to the present fragmented ocean and dispersed continents. Panthalassa had a western prolongation, called Thetys, inserted like a wedge between Eurasia to the north, and Africa, India and Australia to the south. Its margin was almost entirely a consuming plate boundary where oceanic lithosphere was being subducted under the continent. Very little is known about its palaeo-oceanography as most, if not all, of the Triassic (~ 200 Myr) lithosphere has since returned into the mantle. Future drilling may, however, obtain Lower Jurassic sediments from the oldest portion of the Pacific ocean, which is now in the process of disappearing along the western consuming plate boundary, thus providing essential data on Panthalassa.

There is much better information on the evolution of the world ocean during Mesozoic, 165–60 Myr ago, which was characterised by ocean basins in an early stage of evolution, the newly born Atlantic and Indian basins, by a mature but shrinking proto-Pacific and a disappearing Tethys. The breaking up of Pangea resulted in the creation of rifted inactive margins where a great proportion of the Earth's sediments have now been trapped, bordering shallow (≈ 3 km) and narrow basins underlain by hot oceanic lithosphere. The ocean environment under an equable climate, was fairly tranquil with thermo-haline homogeneity and mostly latitudinal, probably wind-controlled circulation pattern (for example, ref. 7).

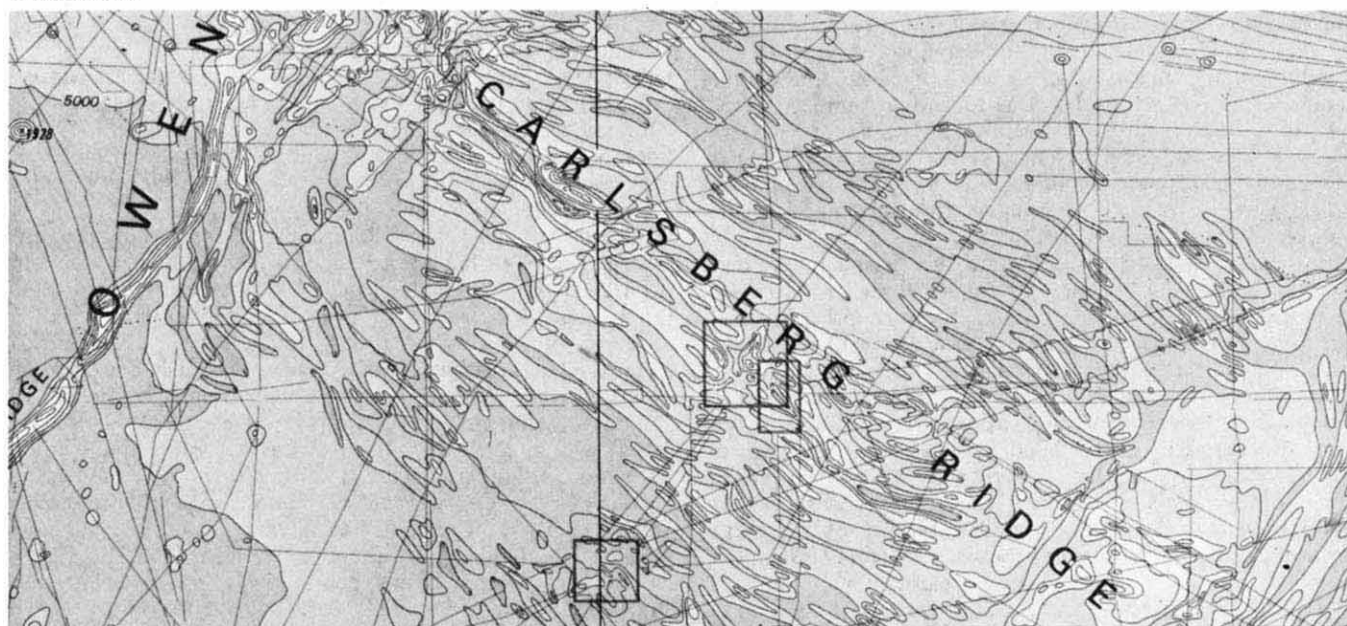
The new ocean basins continuously widened but also deepened, as they became older. The relationship between ocean-depth and age of the plate is a direct consequence of the lithospheric plate being a thermal boundary layer which moves away from its zone of creation at the axis of the mid-ocean ridge. It is a powerful tool to reconstruct the palaeobathymetry of a given ocean, knowing its age.

This age–depth relationship also implies that the average depth of the world ocean is related to its average age. Thus a faster rate of lithosphere renewal should lead to overflow over the continental borderlands, and vice versa. This effect may be the principal cause of the worldwide transgressions and regressions, outside of the geologically relatively recent changes in the volume of the ice-caps. The well known Upper Cretaceous (80 Myr ago) transgression, during which 85% of the Earth was covered by water (compared with 71% today and 65% during the largest regressions), has been explained in this way⁸. In any case,

Then.....



.....and now



A General Bathymetric Chart of the Oceans was first published in 1903 by Prince Albert of Monaco. The upper picture shows a portion of the 1938 edition for the Indian Ocean just east of Somalia (1,400 × 900 km). This chart was still in use at the beginning of the International Indian Ocean Expedition in 1959. The 1975 edition (lower) shows the same area as now charted. The Carlsberg Ridge is now seen to run into the Owen Fracture Zone. Contours are at 1,000 m spacings; ship's tracks are marked lightly.

recent oceanographic, mostly oil-applied research, has demonstrated the existence on the continental margins of numerous global unconformities which are related to sea-level changes. It now seems possible to reconstruct a fairly precise history of the evolution of sea level during the last 200 Myr.

The radical change brought by the Cainozoic (<60 Myr) era, which transformed the warm ocean and mild climates

of the Mesozoic into the cold ocean and glacial climates of the Quaternary (<3 Myr) has been described as a 'commotion in the ocean'. The most dramatic change occurred 38 Myr ago with a drop of about 10 °C of the deep water temperature. This was related to the formation of a circumantarctic seaway, with separation of Australia from Antarctica and the opening of the Drake passage, and to the closure of the Tethys. Strong thermo-haline deep

circulation became well developed, resulting in widespread sea-floor erosion and the present meridional as well as latitudinal circulation pattern was established at this time.

But, it is still not understood why, whereas the first sea-ice appeared 38 Myr ago, the Antarctic ice-sheet did not develop for another 20 Myr. As for the Arctic ice-sheet, it started to develop 3 Myr ago and this event may be related to the creation of the Central America land bridge. The history of the repeated Late Pliocene–Pleistocene (<3 Myr) southward transgression of polar water by as much as 20° of latitude is now being very precisely reconstructed. One can expect in the years to come major progress in the understanding of this Cainozoic 'commotion' from such ocean sediment studies and from oceanologic and climatic modelling.

Geochemistry and geology

The role of the world ocean as a geochemical sink is also being progressively revealed. The degree of oxygenation of the deep ocean has greatly changed. It seems to have been mostly poorly oxygenated during Mesozoic whereas it is highly oxygenated nowadays. These changes probably deeply affected the carbon cycle in the hydrosphere and the carbon dioxide contents of the atmosphere. The development of calcareous and siliceous microorganisms during the Mesozoic must also have introduced a basic change in this geochemical complex cycle.

In contrast, there has been relatively little progress in our understanding of the geological structures generated or modified by the processes of plate tectonics, and first the oceanic lithosphere itself. A striking example concerns its upper portion: the oceanic crust. We know its age nearly everywhere but we still do not know unambiguously its structure and composition. Fifteen years ago, marine geophysicists had proposed, on the basis of seismic refraction results, that the oceanic crust had a homogeneous and uniform three-layer stratification (1, sediments; 2, basement; 3, 'oceanic layer') on top of the upper mantle. Layer 2 was supposed to be made of basalt and layer 3 of gabbro. Hess later proposed that layer 3 consisted instead of serpentinised peridotite. Today, the controversy is still not resolved. The oceanic crust models proposed nowadays are partly based on the results of scattered, poorly located dredgings out of a precise structural frame work but they mostly draw on analogies with ophiolites and predictions based on highly simplified theoretical models.

To this day, drilling has failed to shed much light on this problem because of technological limitations. The largest penetration in the igneous crust does not exceed 600 m,

that is about one tenth of the thickness of the crust. In addition, the necessity to dispose of a least 100 m of sediments over the igneous crust to stabilise the drill bit excludes most of the actively accreting boundaries from being sampled. Consequently, no direct information can be obtained on the actual processes of formation of the crust by drilling. It is fair to conclude that our poor knowledge of the oceanic crust results from the scarcity of direct geological evidence.

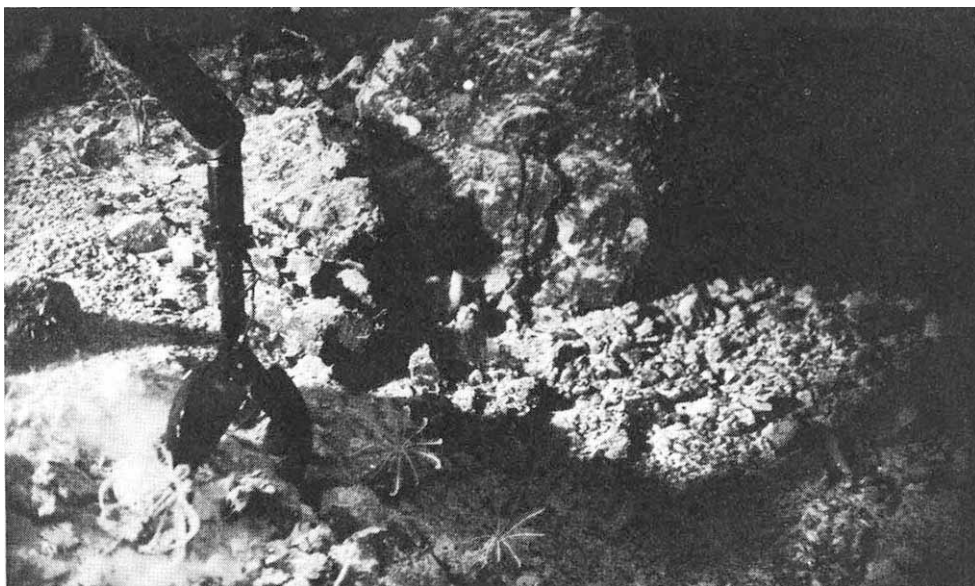
An important development, in this domain, has been the introduction of a much more refined seismic refraction investigation of the ocean crust and lithosphere, with ocean bottom seismographs (the so-called OBS) and better inversion techniques. The resulting acoustic structure is apparently quite complex and variable, as would be expected. Layer 2 seems to be a zone of strong velocity gradient with an uppermost very low velocity zone which progressively disappears with age. There is no generally accepted standard structure for layer 3 and its evolution with age, although it has recently been proposed that a rapid thickening away from the rise crest is due to the formation of a low velocity zone at its base. There is some fairly strong evidence for the existence of a magma chamber at a depth of 2 km below the sea floor under the East Pacific rise¹⁰. No such magma chamber could be detected under the crest of the slower spreading mid-Atlantic ridge.

Very long seismic refraction lines and earthquake seismological studies (primarily surface waves) have given some important information on the structure and evolution of the lithosphere with age. If the thickness of the lithosphere is defined by the lid of the underlying low velocity zone, there is a definite thickening with age from less than 50 km to 80–90 km. In the accretion zone, under the crest of the mid-ocean ridge, where the lithosphere is being created, a zone of partial melting comes within a small distance from the sea floor. Strong internal stratification seems to be common within the lithosphere.

Although this whole field of direct modern investigation of the structure and evolution of the crust and lithosphere by explosion and earthquake seismology is very promising, it is still in its infancy and it is too early to make firm conclusions. Note that an interesting new technique may well be the use of commercial multi-channel seismic reflection for recording oblique reflections from the crust and mantle. Another important new technique will certainly be the use of seismic receivers in the deep-sea drilled holes.

In the absence of detailed direct measurements, physical models based on plate tectonic concepts could bridge the

Mechanical arm of CNEXO submersible CYANA sampling a piece of pillow at 2,700 m depth in Transform Fault A in the Atlantic Ocean during the FAMOUS project. Part of a vertical dyke is seen in the background and numerous signs of animal life.



gap in our knowledge. It is true that it is relatively easy to explain heat flow and water depth variation as a function of age (see ref. 11, for example). But, this success does not mean that the formation of the oceanic crust and lithosphere is understood. Most of the proposed models are purely thermal and ignore the mechanical definition of the lithosphere. Further, widely different sets of parameters can be chosen, provided that the principle of a rigid thermal boundary layer is preserved. And, most important, their results away from the axis do not depend on the mode of transfer of material from the asthenosphere to the lithosphere in the rather narrow accreting zone. Yet, the knowledge of this mode of transfer is essential to any modellisation of the formation of the lithosphere and crust. Clearly, additional constraints are needed. An interesting new one has recently been shown to be that the ocean depth does not increase as fast beyond 60–80 Myr as would be required if it were the cooling of a thermal boundary layer, which indicates the existence of additional sources of heat.

Thus, our knowledge of the ocean crust and lithosphere structure is still schematic. Is there a clearer understanding of the geological structures created by the interaction of plates along their boundaries? In fact, most of the geological models proposed for a consuming plate boundary (deep-sea trench), for an accreting plate boundary (spreading centre), or for a transform boundary are in great part hypothetical and grossly simplified.

Yet, approximately 120,000 km of length of plate boundaries lie below 2.5–11 km of water, which have no equivalent or directly comparable tectonic features above sea level. This is about 80% of the total length of the boundaries of the well defined plates. It is along the oceanic accreting boundaries that the tectonic fabric of the ocean floor is created. It is along the oceanic consuming boundaries that some of the tectonic fabric of the future subaerial mountain belts is presumably set into place. An understanding of the process of formation of the oceanic crust requires knowledge of the spatial and temporal distribution of the volcanic and tectonic events leading in turn to a full comprehension of the pattern of geochemical variation within the accreting plate boundary area. An understanding of the process of subduction also requires a knowledge of the spatial and temporal distribution of the tectonic features within the consuming plate boundary area and this, in turn, can lead to important information on the rheology of the plate itself. Clearly, any coherent development of the plate tectonic model passes by an adequate study of these boundaries.

FAMOUS

Standard oceanographic techniques currently used by most investigators are not adequate for this type of detailed geological and geophysical surveys as they have a spatial resolution which is not much better than half of the water depth. As a result, so-called detailed surveys often have a grid spacing of 10 km or more. Yet, studies of magnetic anomalies using precisely navigated deep-tow studies, have shown that the width of the zone of creation of oceanic crust in accreting plate boundaries is 2–6 km. To resolve processes occurring on such a scale, one needs to use techniques having a much better resolution with very precise navigation. These techniques should of course include bathymetry, seismic refraction survey and detailed geological surface mapping with precisely located multiple drillings. Such a complete study has not yet been performed. The FAMOUS program has made an important step in this direction^{12–13}, however.

The French American Mid-Ocean Undersea Study (FAMOUS) was designed 'to conduct an integrated and detailed geological, geophysical and geochemical study of a portion of plate boundary where new oceanic crust is generated' in the North mid-Atlantic ridge. The main

characteristic of this study is that, for the first time in deep ocean, object having a size from a few centimetres to kilometres were mapped and placed within the framework of a precise structural context. This was done in a first phase with the help of precisely navigated surface ship studies using the powerful multi-narrow beam echo-sounder, deep-tow fish and transponder located dredges and bottom photographs. Then, in a second phase, precisely navigated manned research submersibles established well documented geological sections, with both tectonic observations and rock sampling. Thus, methods of classical field geology could be applied to the deep ocean floor. As a result, a precise structural and petrological model of the Rift Valley and intersecting transform fault has been obtained for the first time.

But FAMOUS was just a beginning. There is little doubt that future research will rely more and more on studies with the same amount of resolving power, from seismic refraction to field geology. A good example is the problem of hydrothermal circulation in newly formed oceanic crust which has been shown to have a major role in the cooling of the crust, the distribution of heat flow, the formation of mineral deposits leached from the crust and the general geochemical cycle of seawater. The vents through which hot water comes to the sea-floor have such a small size that they require high-resolution techniques to be studied. The research submersible Alvin is now making such a study on the Galapagos rise, in the Eastern Equatorial Pacific ocean.

Perhaps, because of the formidable depths at which they lie, consuming plate boundaries have been studied in less detail¹⁴. However, a great step forward in their exploration was made with the acquisition of good commercial seismic reflexion lines revealing the structure of the inner wall which actually consists of a sedimentary accretionary wedge, made of imbricated thrust sheets which is several tens of kilometres long and up to tens of kilometres thick. The actual process of formation of this complex wedge is still not well understood and may include alternating phases of erosion and accretion. It has even been proposed that considerable erosion of continental crust may occur. The importance of such processes in an understanding of the geodynamics cannot be over emphasised.

The future

The next 2 yr, which will see the first systematic attempts at drilling consuming plate boundaries, should shed considerable light on these problems. But a much greater effort is needed to resolve their detailed structure and tectonics. For example, the detailed pattern of microseismicity across a trench is still essentially unknown, although a few OBS measurements have been made.

It is quite probable, however, that, in the next 4–5 yr, the greatest effort of the marine geoscience community will go towards solving the structure and evolution of passive continental margins which contain, as mentioned above, up to 50% of all the Earth's sediments. In addition to their potential economic interest, the passive continental margins are probably the only places where one can find the record of the early stages in the formation of a new ocean. There is still considerable disagreement on such fundamental problems as the nature of the continental crust—ocean crust transition, the actual boundary of the original continent, the processes which have allowed the thinning of the continental crust and the relative importance of the possible processes of subsidence. The ultimate tool for this type of investigation, very deep drilling, is actually being actively considered in spite of the very large costs involved.

For this review to have been truly complete, many other results as well as planned investigations would have had to be mentioned. In addition, no single writer can possibly cover adequately such a large field, so this review is certainly somewhat biased. There is, however, little doubt that the large scale exploration phase of the world ocean

made in the framework of plate tectonics in the last 10 yr is now coming to an end. The next phase will require much more elaborate, more costly techniques, having a greater resolving power to lead to a correct modelling of the various physical and chemical processes involved in the functioning of the great lithosphere engine.

¹ Morgan, W. J. *J. geophys. Res.* 73, 1959–1982, 1968.

² Hess, H. H. in *Petrologic studies—Buddington Memorial vol. Geol. Soc. Am., New York, N.Y.*, 599–6620 (1962).

³ Vine, F. J. & Matthews, D. H. *Nature* 199, 947–949 (1963).

⁴ Wilson, J. T. *Nature* 207, 343–347 (1965).

⁵ *Initial Reports Deep Sea Drilling Project I-XXXVII*, (US Government Printing Office, Washington, 1969–1977).

⁶ Donn, W. L. & Shaw, D. M. *Bull. geol. Soc. Am.* 88, 390–396 (1977).

⁷ Berggren, W. A. & Hollister, C. D. *Tectonophysics* 38, 11–48 (1977).

⁸ Hays, J. D. & Pitman, W. C. *Nature* 246, 18–22 (1973).

⁹ Shackleton, N. J. & Kennett, J. P. *Initial Reports Deep Sea Drilling Project XXIX*, 801–807 (U.S. Government Printing Office, Washington, 1975).

¹⁰ Rosendahl, B. R. *J. geophys. Res.* 81, 5300–5314, 1976.

¹¹ Sclater, J. G. & Francheteau, J. *Geophys. J.* 20, 509–542 (1970).

¹² Arcyana, X. Y. *Science* 190, 108–116 (1975).

¹³ Ballard, R. D. *et al. Science* 190, 103–108 (1975).

¹⁴ Burk, C. A. & Drake, C. L. *The Geology of Continental Margins*, (Springer, New York, 1974).

Breaking waves

E. D. Cokelet

Institute of Oceanographic Sciences, Wormley, Godalming, Surrey, UK

Why do waves break? What happens during breaking itself? More than a hundred years of mathematical studies still leave some aspects of these questions unanswered, although we are now understanding much about the processes leading up to breaking and the period after breaking has occurred.

BREAKING waves, both whitecaps and surf, are commonly observed in the ocean, but surprisingly little is known about them. Yet they are important for a variety of reasons. They place large hydrodynamic loads on man-made structures, transfer horizontal momentum to surface currents, provide a source of turbulent energy to mix the upper layers of the ocean, move sediment in shallow water, and enhance the air-sea exchange of gases and particulate matter. To further our understanding of these processes we must first understand the dynamics of wave breaking.

The scientific study of breaking waves is difficult for a variety of reasons. In theoretical investigations the equations of motion are nonlinear, and the boundary conditions (also nonlinear) must be applied at the free surface whose location is not known beforehand but must be determined as part of the solution. Observations are hindered because breaking occurs rapidly and intermittently. In the ocean large wave forces tend to carry away the instruments; whereas measurements made in controlled laboratory conditions cannot always be scaled-up owing to the relatively greater importance of viscosity and surface tension at the small scale.

At the present stage of theoretical development simplifying assumptions are commonly made about the flow-field in a wave. The dynamics of ocean waves are dominated by the inertia of the fluid, the pressure gradient, and the acceleration of gravity. For large scale waves the fluid viscosity can be neglected in the bulk of the fluid since it is important only in very thin layers near the fluid boundaries. Similarly, surface tension usually can be neglected unless the free surface is very highly curved. It is mathematically convenient to assume the fluid to be irrotational, that is, to have no vorticity—a measure of the rotation which an individual element of fluid has about its centre. This will be true if a wave propagates into a region of inviscid fluid previously at rest. It will always be a very good approximation unless there is a strong current shear or the flow becomes turbulent. In the latter case the fluid can still be irrotational outside the turbulent zone. The mathematical simplification is considerable since in an irrotational wave the vector velocity may be expressed as the gradient of a scalar (the velocity potential function) for which a linear equation (Laplace's equation) holds. We can neglect the airflow above the water since the density of air is 10^{-3} times that of water. Finally it is useful to consider the flow to be two-dimensional, that is, to neglect variations in a direction parallel to the wave crest. As more is

learned about the wave motion the effects of these neglected quantities can be incorporated.

It will be helpful to review some of the properties of simple waves which do not break. A wave is characterised by its wavelength (λ), water depth (d) and wave height (H , see Fig. 1). It is not the absolute values of these quantities but their ratios which are important. When the water depth is greater than about half of a wavelength the wave is in deep water, and it does not feel the effects of the bottom. Smaller values of d/λ correspond to shallow-water waves, and in the limit when d/λ approaches 0 a solitary wave exists, so-called because it consists of a solitary crest moving in shallow water with troughs infinitely far away. Many of the simple properties of non-breaking ocean waves can be explained in terms of linear, small-amplitude theory which is an approximation to the complete theory of wave motion valid when the wave steepness (H/λ) and height-to-depth ratio (H/d) are much less than 1. In deep water when a small-amplitude wave passes, individual particles of water do not travel along with the wave, but instead they move in closed circular orbits whose radii decrease exponentially with depth as shown in Fig. 1. In shallow water these trajectories become ellipses with shortened vertical axes, and for a solitary wave each fluid particle moves forward with a hop as the wave passes. The profile of a small-amplitude wave is symmetric about the vertical line passing through its crest.

Research work on breaking waves can be divided into three categories: those concerning waves (1) before, (2) during, and (3) after breaking. In this review we shall concentrate on some of the latest results in the first two categories.

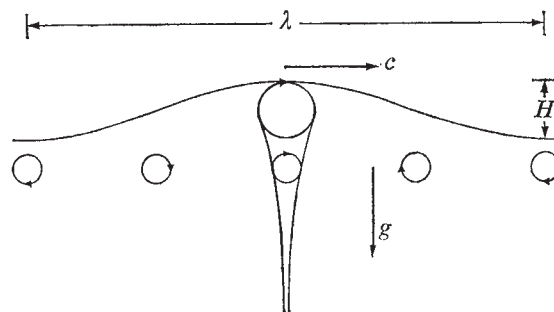


Fig. 1 The profile of a small-amplitude, sinusoidal wave in deep water showing the circular orbits of fluid particles with arrowheads indicating the direction of water movement at various locations. The rapid decrease of the orbital radii with depth is illustrated beneath the crest. The wave of wavelength λ (the horizontal distance between successive troughs or crests) and height H (the vertical crest-to-trough distance) is moving to the right at speed c with gravity g providing the external restoring force. (This wave is too high relative to its wavelength for small-amplitude theory to hold, but it was drawn this way for purposes of illustration).

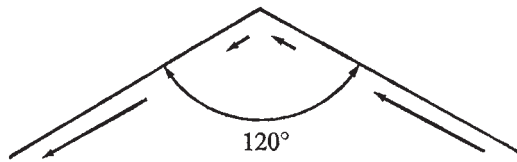


Fig. 2 A close-up view of the crest of the highest, steady wave as seen from a reference frame moving with the wave to the right. Fluid particles come to rest at the sharp wave crest.

Before breaking

One approach to the study of breaking waves has been to start with a small-amplitude wave and to consider progressively higher symmetric waves which propagate without change of shape in water of constant depth. For a fixed wavelength and depth there exists a highest wave which is just on the verge of breaking. When viewed from a reference frame moving with the wave the flow is steady, that is, independent of time, and this simplifies the problem. Analytically it is possible to begin with the small-amplitude approximation to the fluid motion and to correct this for the effects of increased wave height. This is then recorrected and so on until sufficient accuracy is obtained. Each successive correction may be thought of as one member in an ordered series of corrections called a perturbation series. The importance of any correction relative to the one before it is measured by a perturbation parameter which is related to the wave height. To get an accurate answer for higher waves the perturbation series must be taken to higher order, that is, more and more corrections are needed. Sir George Stokes found the third-order correction in 1847 (ref. 1) for periodic waves in water of arbitrary, uniform depth and the fifth-order, deep-water correction in 1880 (ref. 2). His results indicate that for low-but-finite-amplitude waves the speed of wave propagation (the phase speed) increases with the wave height. The fluid particles no longer move in closed orbits, but they acquire a small, added, down-wave drift called the Stokes drift. Also the wave profile does not remain symmetric about its mean level as in a small-amplitude sinusoidal wave, but instead the crests become higher and sharper and the troughs flatter. Unfortunately it is not possible to carry out the Stokes expansion to a high enough order by hand to accurately determine the properties of the highest wave because the amount of algebraic manipulation required increases dramatically with order. Indeed the tenth-order, deep-water correction was not obtained until 1914 (ref. 3), and the fifth-order, general-depth one was found in 1955 (ref. 4). Slow progress has also been made for solitary waves⁵⁻¹² with the third-order result being obtained in 1971 (ref. 12).

A separate and completely independent line of reasoning yields some information about the highest wave. In general at a wave crest the fluid particles are moving forward at a speed less than the phase speed (c), and as higher waves are considered the particle velocity approaches the phase speed. In a limiting wave these two velocities are equal. A further increase in wave height will cause the fluid particles to overtake the wave itself, and breaking will ensue. In the reference frame moving with the wave the water comes to rest at the crest, and it must attain a sharp point¹³ with an included angle of 120° as shown in Fig. 2. This corner-flow is only a first approximation valid very close to the crest and includes no information about the rest of the wave or about lower waves.

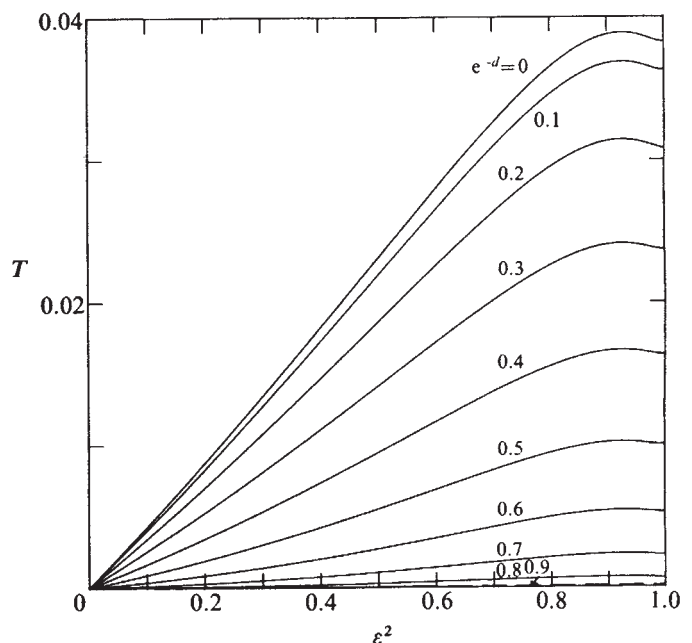
There have been several attempts to bridge the gap between the finite-amplitude theories and the highest wave. Some theoreticians have tried to build the 120° angle into the theories¹⁴⁻¹⁷, but it is not clear how this should be done¹⁸⁻²⁰. Many purely numerical approaches have been tried²¹⁻³⁰, and these are successful for small to moderately-high waves. They usually involve guessing the wave shape and then adjusting this until some error criterion is satisfied at a number of points along the wave profile. The advantage of these methods is that a high-

speed computer can be used to do the calculations, but often they fail to reach acceptable accuracy levels when the highest wave is approached.

Some recent advances have been made with perturbation expansion methods in which the laborious algebraic manipulations are carried out automatically by computer^{19,31-34}. The series can be expanded to a very high order (past order 100 in some cases) and summed efficiently thus providing accurate answers up to very near the highest wave. Stokes' original expansion procedure fails for high waves since his expansion parameter does not always increase with the wave height¹⁹; therefore two distinct waves exist for the same value of this parameter. The wave height itself can be used as an expansion parameter¹⁹, but its maximum value is not known beforehand. This difficulty is overcome by defining a new parameter (ϵ) in terms of the fluid speeds at the wave crest (q_{crest}) and trough (q_{trough}) in the moving reference frame³²⁻³⁴. The complete range of this parameter ($\epsilon^2 = 1 - q_{\text{crest}}^2 q_{\text{trough}}^2 / c^4$) is known from the beginning. For a small-amplitude wave the water streams past to the left at the phase speed ($q_{\text{crest}} = q_{\text{trough}} = c$, $\epsilon = 0$), and for a highest wave it comes to rest at the crest ($q_{\text{crest}} = 0$, $\epsilon = 1$). The wave height increases steadily with ϵ to values of $H/\lambda \approx 1/7$ for deep-water waves and $H/d \approx 5/6$ for solitary waves, but other wave properties behave in a surprising manner. The phase speed, momentum and energy increase with wave height initially, reach maxima, and then decrease as the highest wave is approached³²⁻³⁴. Therefore the highest wave is not the most energetic as had always been assumed. This behaviour is illustrated in Fig. 3 which is a typical plot of the kinetic energy of the wave against ϵ^2 for various ratios of water depth to wavelength³⁴.

The energy maxima can be explained in terms of the wave profile^{32,34,35}. At the crest the profile of the highest wave is above that of a slightly lower wave, but because of the sharp peak it falls away more rapidly and soon dips below the lower wave. Considered as an average over a wavelength, the highest wave is actually lower than an almost-highest wave; hence

Fig. 3 The average kinetic energy per unit surface area (T) of a steady wave plotted against the expansion parameter, $\epsilon^2 = 1 - q_{\text{crest}}^2 q_{\text{trough}}^2 / c^4$, for various water depths (d) (ref. 34). For each depth the highest wave corresponds to $\epsilon^2 = 1$. Notice that each curve reaches maximum energy for waves lower than the highest. Quantities are measured in units of density (ρ), wave-number ($k = 2\pi/\lambda$) and gravity (g).



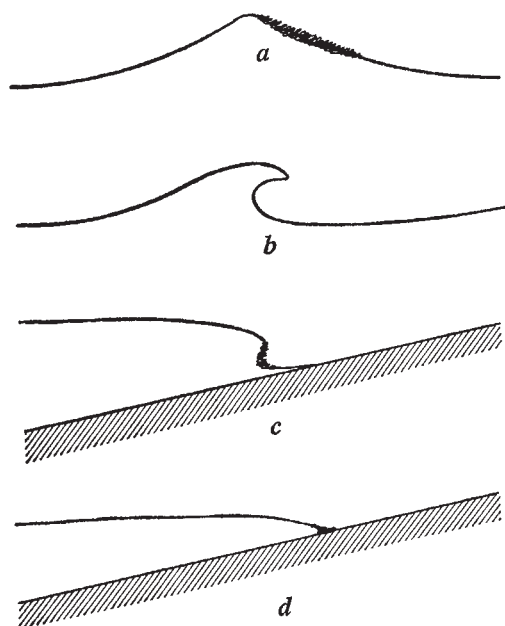


Fig. 4 The four main breaker types: *a*, spilling; *b* plunging; *c*, collapsing, and *d*, surging.

it has less energy. These surprising results have been confirmed by separate numerical³⁵ and analytical³⁶ approaches which are completely independent of the perturbation series.

The existence of the maxima in the wave property curves may have a bearing on why high symmetric waves begin to break^{32,34}. Usually a wave is specified uniquely by the water depth, wavelength and one other wave property. But because the curves of these properties bend over for very high waves two distinct states are possible with the same wave energy, say, but with different wave heights. The solutions corresponding to one part of the curve may be unstable, and a wave might jump from one part to the other. Alternatively symmetric waves may become unsymmetric at the peak of their energy curves making the highest wave inaccessible except through the loss of energy. Other possibilities exist, and there is scope for further research in this area.

Several laboratory experiments have been carried out to test the validity of the steady wave theories. Usually low waves are generated in a long, narrow tank with a paddle, piston or wedge-shaped plunger and are sent into a channel with a sloping bottom or converging side walls to build up their height. In the test section simple measurements of the water depth, wave height and phase speed can be made. The measured phase speeds and limiting wave heights agree with theory to within experimental error (a few per cent), but the observations are not precise enough to detect the phase speed maxima³⁷⁻⁴¹. Indeed it is difficult to produce very-high waves which remain steady and do not break. Whether this is due to the method of wave generation or due to an inherent instability has yet to be determined. In more complex experiments the surface profile, velocity distribution, particle orbits and accelerations have been measured⁴¹⁻⁴⁵. Comparisons with low-order, finite-amplitude theories (Stokes' third order) in deep water are acceptable, but they worsen as H/λ and H/d increase. Differences of 50% between the theoretical and the measured (theoretical > measured) fluid velocities are common for $d/\lambda = 1/10$ and wave heights of about half the maximum. This is unacceptable for design purposes since the hydrodynamic drag on a structure varies as the square of the velocity so that a 50% velocity error produces a 100% drag error. The most widely used numerical method²³ does somewhat better, but it is inaccurate for high waves in all water depths³⁴. The discrepancy between theory and observation is probably due to the low order of the theories used. Their slow convergence rate for

shallow-water is well known^{1,2}, and it is quite likely that the inclusion of higher-order terms will bring about a marked improvement. The results of the recent high-order calculations^{19,31-34} have yet to be compared in detail with experiments.

Surface tension becomes increasingly important as the curvature of the wave crest increases, and this effect will be amplified in a small-scale laboratory experiment where the wavelength is 10^{-2} that of a typical ocean wave. Surface tension makes it impossible for the highest wave crest to attain a sharp point, and what this will do to the energy maxima is not known. Ripples are generated at the highly curved crest, and these ride along on the forward face of the wave^{46,47}. At present no high-order theories include both gravity and surface tension, but this is a future goal. An exact theory does exist for pure capillary waves in which surface tension provides the restoring force and gravity is neglected⁴⁸⁻⁴⁹, but owing to viscosity this is hardly, is ever, applicable.

During breaking

We now turn to the time-dependent behaviour of waves as they steepen and overturn. Roughly speaking a wave will tend to break when the local energy density (E) exceeds a critical value ($E_{\max}$), that of the most energetic steady wave of the same wavelength and water depth. This happens during shoaling when the energy scale ($E_{\max}$) decreases shoreward, and $E/E_{\max}$ increases until the break point.

Results from laboratory experiments with shoaling waves indicate that breakers can be classified into four main types⁵⁰⁻⁵⁵—spilling, plunging, collapsing and surging—as illustrated in Fig. 4. The breaker type depends upon both the initial wave energy, characterised by the offshore wave steepness ($H_{\infty}/\lambda_{\infty}$), and the rate of energy input due to shoaling, characterised by the beach slope (s) (refs 40, 52-54, 56). A breaker parameter (B) representing the ratio of these two effects can be defined by $B = H_{\infty}/(\lambda_{\infty}s^2)$ (refs 53, 54). As B decreases from large values there is a progression from spilling to surging breakers. A spilling breaker which occurs when a steep wave approaches a gradually shoaling beach is similar in profile to a steep, symmetric wave, but it has a turbulent plume of frothy water running down its forward face (Fig. 5). Apart from this turbulent region the flow is irrotational. A value of B of approximately 5 marks the transition from spilling to plunging. In a plunging breaker the wave steepens and overturns ejecting a jet of water from near its crest which eventually strikes the surface below (Fig. 6). The flow may remain laminar (non-turbulent) until right up to the point of impact, or the jet may become turbulent due to the growth of small scale instabilities induced by the wind and surface tension. The third breaker type is the collapsing breaker in which the wave overturns not at its crest but lower down the forward face, and a region of white water

Fig. 5 A spilling breaker in shallow water (photo copyright Institute of Oceanographic Sciences).





Fig. 6 A plunging breaker in shallow water (photo copyright J. D. Humphrey).

forms probably due to the close proximity of the bottom. This type is not as clearly defined as the others, and it denotes the transition between plunging and surging at $B \approx 0.1$. The surging breaker moves up a steep beach as the forward tip of a low wave. It resembles a standing wave formed by total reflection from a vertical wall, but a small amount of turbulence may be present at the beach face.

Two of the shallow-water breaker types—spilling (Fig. 7) and plunging—can occur in deep water. The bottom slope is no longer relevant since the bottom is far away and does not affect the dynamics. Deep-water waves can receive an input of energy when they overtake or collide with one another, meet an opposing current or receive a 'push' from the wind. In a laboratory wave tank the energy input can be provided by a converging side wall which acts in a way similar to a sloping beach. There is a transition from spilling to plunging as the initial wave steepness decreases and the channel convergence increases⁴¹. For gently spilling waves $H/\lambda \approx 1/7$ which confirms that a spilling breaker may be thought of as a just-breaking, highest, symmetric wave. Collapsing and surging breakers do not occur in deep water, but their counterparts are a mixture of progressive and standing waves. In the limit when two identical wave trains propagating in opposite directions collide there exist pure standing waves. The highest standing wave of periodically recurring permanent shape has $H/\lambda \approx 2/9$ (refs 57, 58). Not much is yet known about the breaking characteristics of standing waves or of mixed waves.

Experiments with breaking waves in shallow water have concentrated on how the ratio of wave height to water depth at the break point (H_b/d_b) and the ratio of breaking wave height to initial wave height (H_b/H_∞) vary with initial wave steepness and beach slope. Some empirical relationships based on scattered experimental data have been derived⁵⁹, but not a great deal of confidence can be placed in them. Generally for spilling breakers H_b/d_b is in the range 0.65–0.85 (refs 50, 52, 54, 60) which is approximately that value obtained from the highest solitary wave theory. For plunging breakers H_b/d_b increases, and values as large as 1.3 and 3.0 have been measured for periodic^{50,52} and solitary^{60,61} waves respectively, H_b/H_∞ equals about 1 for gently spilling waves which break almost immediately upon shoaling and increases to about 2 for plunging breakers^{50,52}. The behaviour as the ratio of wave steepness to beach slope tends to zero is not well defined. If the slope is small but finite and the initial wave steepness tends to zero an incoming wave splits up into a series of solitary-like waves (solitons) of different heights, and the highest breaks first as a spilling breaker⁵⁴. But if the initial wave slope is held fixed and the beach becomes a vertical wall (s_∞), an incoming wave will be totally reflected if low enough or break as a surging breaker.

A few measurements of the velocity distribution in spilling and plunging breakers have been made in the laboratory^{41,42,50,52,62,63} and the field⁶⁴. Generally with increasing profile asymmetry the rapid-forward-velocity region beneath the crest deepens, and the velocity vectors there become tilted upward as in accordance with the formation of a plunging jet. On sloping beaches the backwash of the previous wave augments the velocity asymmetry and helps to topple the wave^{50,52,64}. In the most precise experiment⁴¹ the crest velocity of a deep-water plunging breaker was 1.6 times the phase speed just after the onset of breaking.

No analytical theories adequately describe the time-dependent process of wave breaking. Theories exist for small-amplitude, shallow-water waves, but these are approximations valid only when the wave slopes are low or when the fluid acceleration is small compared with gravity^{65,66}. When the free surface steepens the fluid acceleration becomes comparable with gravity and cannot be neglected. An attempt has been made to study the effect of a small perturbation on the 120° corner-flow⁶⁷, but this theory is not appropriate since the solution contains undesirable mathematical singularities both before and after the instant of breaking.

Modern numerical techniques have had some success in describing breaking. The method most widely used by ocean engineers involves actually measuring the profile of a breaking wave and then fitting an irrotational flow to this by minimising the departure from constant atmospheric pressure along the surface²³. It has three major disadvantages: (1) the wave shape is assumed to be steady which is certainly not true during breaking, (2) the method cannot represent overturned waves, and (3) it cannot predict the evolution of a wave. It does give a rough estimate of the velocity field, and experimental comparisons indicate velocity errors of 10%–30% beneath spilling breakers at the break point⁶⁸.

A series of finite-difference numerical methods have been devised to deal with free-surface deformations. The two-dimensional fluid region is divided up into a number of small cells, and the velocity and pressure gradients are approximated by their differences across the cells. Marked fluid particles denote the free surface, and these are moved every time step by velocities interpolated from the cell boundaries. An early method⁶⁸ allowed the free surface to curl over, but an artificially large fluid viscosity had to be used to damp out unwanted 'wiggles' (numerical instability). A modified inviscid method⁶⁹ is more accurate and has been used successfully to calculate the run-up of a solitary wave on a steep beach. But it fails at the point of spilling when a solitary wave approaches a gradual beach because it cannot cope with an overturning surface⁷⁰. Similar methods which allow the cells to move with the fluid should do better, but available computer resources limit the minimum cell size so that the cells become too distorted for acceptable accuracy as the wave steepens⁷¹. Attempts to rezone the cells lead to instability at breaking⁷².

A different formulation has made it possible to follow waves as they overturn⁷³. One of the properties of irrotational flow theory is that the shape of the fluid boundary and the value of the velocity potential on that boundary uniquely determine the flow-field inside. Therefore only the evolution of the free surface need be considered, and all the computational effort can be focused there to accurately resolve its shape. At a given instant the tangential and normal velocities of the surface which are needed to follow its evolution can be found by numerical differentiation and by the solution of a linear integral equation. Initially computational particles are distributed on the free surface, and their trajectories are calculated in time steps⁷³. Figure 8a (Cokelet, unpublished work) is a computer-generated plot of the profile of a periodic, deep-water wave at various times as it steepens and overturns. The initial wave moving to the right begins as a sinusoid but with too much energy ($E = 5/3 E_{max}$) to remain symmetric. Figure 8b illustrates a portion of the wave profile near the crest as seen by an observer moving at the speed of a small-amplitude wave. The profiles, plotted from



Fig. 7 A deep-water spilling breaker (or whitecap) moving toward the left (photo copyright J. R. Dunn).

the instant the wave face becomes vertical, show the crest toppling over towards the water below, but the theory does not remain valid past the point of impact. For spilling-like breakers the radius of curvature at the leading edge of the jet approaches the scale of the particle spacing, and the computations must be halted owing to numerical inaccuracies. Since irrotational flow is assumed there is no turbulent spilling, and the neglect of surface tension means no droplets pinch off. Numerical experiments are being conducted to determine how the plunging jet's momentum and energy depend on the initial conditions and energy input.

In some instances contrary to our simplifying assumption the airflow does have an effect on breaking. The wind creates a thin, sheared layer of water at the surface (the wind-drift layer), and this can augment the fluid particle velocity at a wave crest forcing the wave to break at a height below the maximum⁷⁴. Also the wind can blow-off or hold-up the wave crests, but this has not been studied. Conversely, breaking itself can influence the airflow, and this may help to explain how the wind makes waves. The airflow over a wave has a horizontally varying pressure field which must be shifted laterally with respect to the wave to feed energy into it⁷⁵. Theories which assume small-amplitude waves predict a wave growth rate which is much lower than that observed⁷⁵. But some new experimental findings show that breaking forces the airflow to separate from the wave surface thus shifting the pressure field so that it can do more work on the wave⁷⁶. Therefore wave growth is actually triggered by breaking which is a dissipative process.

After breaking

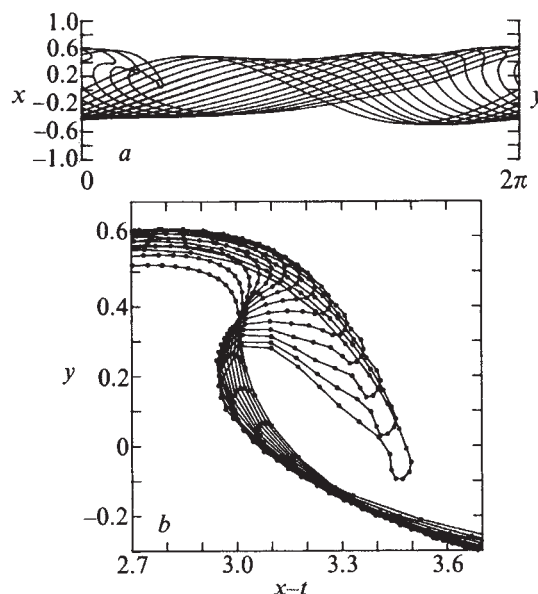
In shallow water after the period of rapid wave steepening and overturning there can exist a much less time-dependent period

in which the energy dissipated in the breaker approximately balances the increase of relative energy ($E/E_{\max}$) due to shoaling.

For a spilling breaker a turbulent, air-water mixture (a plume) runs down its forward face accelerated by gravity and retarded by the entrainment of up-slope momentum from the flow below. Theoretical considerations⁷⁷ based on the assumption of steady flow show that the plume can exist only if the surface slope and density difference between the whitecap and the water are sufficiently large. For a 30° surface slope corresponding to the 120° corner-flow, the plume must be about 8% air which is in agreement with observations of other self-aerated flows. The plume theory indicates that, once formed, a whitecap will tend to spread down the wave face, but energy balance considerations show that the ratio of whitecap length to wave height should be about constant⁷⁷. Since the wave height tends to diminish shoreward of the break point no true steady state can exist. It is observed^{55,77} that the whitecap grows, and its damping effect overcomes the effect of wave steepening due to shoaling. The wave crest becomes rounded which reduces the slope and allows part of the whitecap to be swept back over the crest. The wave then steepens, and the whole process repeats itself as long as sufficient wave energy remains⁷⁷.

After a wave has broken a step-like jump in the free surface can form and advance shoreward as a bore. Due to limitations of space we will not consider bores in detail but instead refer the reader to reviews on them⁷⁸ and on related hydraulic jumps⁷⁹. However, a recent numerical investigation which goes somewhat farther than previous ones should be pointed out⁸⁰. In this the run-up of a bore on a sloping beach is modelled. It is possible to follow the water up the beach until its forward tip gets very thin and then to follow its subsequent rush seaward (backwash) which itself develops into a bore. Preliminary results for a series of bores approaching periodically from the sea indicate that the maximum water velocity actually occurs during the backwash stage which may be of significance for sediment-transport studies.

Fig. 8 a, Computer-drawn profiles of a plunging breaker in deep water at successive time intervals of about $1/25$ of a wave period ($\Delta t = 1/4$ in units of ρ , k and g). The initial wave form is sinusoidal with its crest at $x = \pi/2$, but it has too much energy ($E = 5/3 E_{F\max}$) to remain symmetric. Therefore the wave steepens and overturns. b, A portion of the crest plotted about every $1/50$ of a wave period ($\Delta t = 1/8$) as seen by an observer moving to the right with the speed of a small-amplitude wave ($c^2 = g/k$). The profiles are plotted from when the wave face becomes vertical, and the dots mark the computational particles (Cokelet, unpublished work).



For a tsunami the ratio of offshore wave steepness to beach slope is very small, and the wave usually does not break at all. In that case the maximum shoreward excursion of the water (the run-up distance) is of particular interest to those concerned with tsunami prediction and the design of shoreline structures. Advances have been made in the theory of wave run-up, but it is beyond the scope of this paper to consider them here (see ref. 81 for more details).

Conclusion

We understand some things about breaking waves especially before and after breaking when the flow is steady and more amenable to theoretical and experimental study. But there are large gaps in our knowledge. Much has yet to be learned about the time-dependent period during breaking. Laboratory experiments performed in shallow water have provided some information about this, and numerical investigations are likely to prove fruitful in the future. In this exciting field many questions remain unanswered. What are the velocity and acceleration fields in a breaking wave? How much of its momentum and energy are given up to ocean currents and turbulence? What is the spatial distribution and frequency of breaking? There is great potential for further theoretical and experimental research. The spectacular fury of a breaking wave stimulates the search.

- ¹ Stokes, G. G. *Trans. Camb. Phil. Soc.* **8**, 441–455 (1847).
- ² Stokes, G. G. *Mathematical and Physical Papers I*, 314–326 (Cambridge University Press, Cambridge, 1880).
- ³ Wilton, J. R. *Phil. Mag.* **27**, 385–394 (1914).
- ⁴ De, S. C. *Proc. Camb. Phil. Soc.* **51**, 713–736 (1955).
- ⁵ Boussinesq, J. C. R. *Acad. Sci., Paris*, 755 (1871).
- ⁶ Rayleigh, Lord *Phil. Mag.* **1**, 257–279 (1876).
- ⁷ McCowan, J. *Phil. Mag.* **32**, 45–58 (1891).
- ⁸ Korteweg, D. J. and de Vries, G. *Phil. Mag.* **39**, 422–443 (1895).
- ⁹ Weinstein, A. *Atti Accad. Lincei, Rend. Cl. Sci. Fiz. Mat. Nat.* **3**, 463–468 (1926).
- ¹⁰ Long, R. R. *Tellus* **8**, 460–471 (1956).
- ¹¹ Laitone, E. V. *J. Fluid Mech.* **9**, 430–444 (1960).
- ¹² Grimshaw, R. J. *Fluid Mech.* **46**, 611–622 (1971).
- ¹³ Stokes, G. G. *Mathematical and Physical Papers I*, 225–228 (Cambridge University Press, Cambridge, 1880).
- ¹⁴ Michell, J. H. *Phil. Mag.* **36**, 430–437 (1893).
- ¹⁵ McCowan, J. *Phil. Mag.* **38**, 351–358 (1894).
- ¹⁶ Havelock, T. H. *Proc. R. Soc. A* **95**, 38–51 (1919).
- ¹⁷ Longuet-Higgins, M. S. *Proc. R. Soc. A* **331**, 445–456 (1973).
- ¹⁸ Grant, M. A. *J. Fluid Mech.* **59**, 257–262 (1973).
- ¹⁹ Schwartz, L. W. *J. Fluid Mech.* **62**, 553–578 (1974).
- ²⁰ Norman, A. C. *J. Fluid Mech.* **66**, 261–265 (1974).
- ²¹ Yamada, H. *Rep. Res. Inst. appl. Mech. Kyushu Univ.* **5**, 37–52 (1957).
- ²² Yamada, H. *Rep. Res. Inst. appl. Mech. Kyushu Univ.* **5**, 53–67 (1957).
- ²³ Dean, R. G. *J. geophys. Res.* **70**, 4561–4572 (1965).
- ²⁴ Schwitters, D. J. thesis, Univ. Arizona (1966).
- ²⁵ Thomas, J. W. *Mathematika* **15**, 139–148 (1968).
- ²⁶ Byatt-Smith, J. G. B. *Proc. R. Soc. A* **315**, 405–418 (1970).
- ²⁷ Strelkoff, J. *Proc. 2nd Int. Conf. Num. Meth. Fluid Dyn.* (Springer, New York, 1971).
- ²⁸ Von Schwind, J. J. and Reid, R. O. *J. Geophys. Res.* **77**, 420–433 (1972).
- ²⁹ Sasaki, K. and Murakami, T. *J. Ocean. Soc. Japan* **29**, 94–105 (1973).
- ³⁰ Thomas, J. W. *Q. app. Math.* **32**, 403–410 (1975).
- ³¹ Fenton, J. J. *Fluid Mech.* **53**, 257–271 (1972).
- ³² Longuet-Higgins, M. S. and Fenton, J. D., *Proc. R. Soc. A* **340**, 471–493 (1974).
- ³³ Longuet-Higgins, M. S. *Proc. R. Soc. A* **342**, 157–174 (1975).
- ³⁴ Cokelet, E. D. *Phil. Trans. R. Soc. A* **286**, 183–230 (1977).
- ³⁵ Byatt-Smith, J. G. B. and Longuet-Higgins, M. S. *Proc. R. Soc. A* **350**, 175–189 (1976).
- ³⁶ Longuet-Higgins, M. S. and Fox, M. J. H. *J. Fluid Mech.* **80**, 721–742 (1977).
- ³⁷ Daily, J. W. and Stephen, S. C. *Proc.—Separate No. 107* (American Society of Civil Engineers, 1951).
- ³⁸ Daily, J. W. and Stephen, S. C. *Proc. 3rd Conf. cst. Engng.*, 13–30 (1953).
- ³⁹ Danel, P. *Gravity Waves, U.S. National Bureau of Standards, Circular 521*, 35–38 (US Government Printing Office, Washington D.C., 1952).
- ⁴⁰ Wiegell, R. L. *Oceanographical Engineering* (Prentice-Hall, Englewood Cliffs, 1964).
- ⁴¹ Van Dorn, W. G. and Pazan, S. P. *Scripps Inst. of Ocean., Adv. Ocean Engng. Lab., Rep. AOEL 71* (1975).
- ⁴² Morison, J. R. and Crooke, R. C. *Tech. Mem. No. 40* Beach Erosion Board, US Army Corps Engineers (1953).
- ⁴³ Le Méhauté, B., Divoky, D. and Lin, A. *Proc. 11th Conf. cst. Engng.* **1**, 86–107 (1968).
- ⁴⁴ Iwagaki, Y. and Sakai, T. *Proc. 12th Conf. cst. Engng.* **1**, 309–325 (1970).
- ⁴⁵ Dean, R. G. *Proc. Symp. Long Waves*, 129–152 (University of Delaware, Newark, Delaware, 1972).
- ⁴⁶ Cox, C. S. *J. mar. Res.* **16**, 199–225 (1958).
- ⁴⁷ Longuet-Higgins, M. S. *J. Fluid Mech.* **16**, 138–159 (1963).
- ⁴⁸ Crapper, G. D. *J. Fluid Mech.* **2**, 532–540 (1957).
- ⁴⁹ Kinnorsley, W. J. *Fluid Mech.* **77**, 229–241 (1976).
- ⁵⁰ Iversen, H. W. *Gravity Waves, US National Bureau of Standards, Circular 521*, 9–32 (US Government Printing Office, Washington, DC, 1952).
- ⁵¹ Mason, M. A. *Gravity Waves, US National Bureau of Standards, Circular 521*, 215–220 (US Government Printing Office, Washington, DC, 1952).
- ⁵² Iversen, H. W. *Proc. 3rd Conf. cst. Engng.* **1**–12 (1953).
- ⁵³ Galvin, C. J. *J. geophys. Res.* **73**, 3651–3659 (1968).
- ⁵⁴ Galvin, C. J. *Waves on Beaches* (ed. Meyer, R. E.), 413–456 (Academic, New York, 1972).
- ⁵⁵ Kjeldsen, S. P. and Olsen, G. B. *Breaking Waves* (16 mm film) Coastal Engineering Laboratory, Technical University of Denmark (1968).
- ⁵⁶ Patrick, D. A. and Wiegell, R. L. *Proc. 1st Conf. Ships Waves*, 397–422 (1955).
- ⁵⁷ Penney, W. G. and Price, A. T. *Phil. Trans. R. Soc. A* **244**, 254–284 (1952).
- ⁵⁸ Taylor, G. I. *Proc. R. Soc. A* **218**, 44–59 (1953).
- ⁵⁹ Weggel, J. R. *J. Watways. Harb. Div. Am. Soc. civ. Engrs.* **98**, 529–548 (1972).
- ⁶⁰ Ippen, A. T. and Kullen, G. *MIT Hydrodynamics Lab., Tech. Rep. No. 15* (1955).
- ⁶¹ Camfield, F. E. and Street, R. L. *J. Watways. Harb. Div. Am. Soc. civ. Engrs.* **95**, 1–22 (1969).
- ⁶² Divoky, D., Le Méhauté, B. and Lin, A. *J. Geophys. Res.* **75**, 1681–1692 (1970).
- ⁶³ Iwagaki, Y. and Sakai, T. *Mem. Faculty Engng Kyoto Univ.* **38**, 11–20 (1976).
- ⁶⁴ Miller, R. L. and Zeigler, J. M. *Proc. 9th Conf. cst. Engng* 103–122 (1964).
- ⁶⁵ Airy, G. B. *Encyclopedia Metropolitana* London (1845).
- ⁶⁶ Biesel, F. *Gravity Waves, US National Bureau of Standards, Circular 521*, 243–253 (US Government Printing Office, Washington, D.C., 1952).
- ⁶⁷ Price, R. K. *J. geophys. Res.* **76**, 1576–1581 (1971).
- ⁶⁸ Harlow, F. H., Shannon, J. P. and Welch, J. E. *Science* **149**, 1092–1093 (1965).
- ⁶⁹ Chan, R. K.-C. and Street, R. L. *J. comp. Phys.* **6**, 68–94 (1970).
- ⁷⁰ Chan, R. K.-C. and Street, R. L. *Proc. 12th Conf. cst. Engng* **1**, 345–361 (1970).
- ⁷¹ Brennan, C. and Whitney, A. K. *Proc. 8th Symp. Naval Hydro.* 117–145 (Office of Naval Research, Washington, DC, 1970).
- ⁷² Chan, R. K.-C. *J. comp. Phys.* **17**, 311–331 (1975).
- ⁷³ Longuet-Higgins, M. S. and Cokelet, E. D. *Proc. R. Soc. A* **350**, 1–26 (1976).
- ⁷⁴ Banner, M. L. and Phillips, O. M. *J. Fluid Mech.* **65**, 647–656 (1974).
- ⁷⁵ Barnett, T. P. and Kenyon, W. E. *Rep. Prog. Phys.* **38**, 667–729 (1975).
- ⁷⁶ Banner, M. L. and Melville, W. K. *J. Fluid Mech.* **77**, 825–842 (1976).
- ⁷⁷ Longuet-Higgins, M. S. and Turner, J. S. *J. Fluid Mech.* **63**, 1–20 (1974).
- ⁷⁸ Rajaratnam, N. *Adv. Hydrosci.* **4**, 197–280 (1967).
- ⁷⁹ Witting, J. *J. CRC Critical Reviews in Solid State Sciences*, **2**, 555–581 (1972).
- ⁸⁰ Hibberd, S. and Peregrine, D. H. *Proc. IUTAM sym. on waves on water of varying depth*, 20–23 July 1976. (Australian Acad. Sci., Canberra, 1977).
- ⁸¹ Meyer, R. E. and Taylor, A. D. *Waves on Beaches* (ed. Meyer, R. E.), 357–411 (Academic, New York, 1972).

Acoustic oceanography

Walter H. Munk & Gordon O. Williams

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093

Measurements of ocean finestructure (with typical time scales of 15 min to 1 d, spacial scales of 10 to 100 m vertically and 1 to 10 km horizontally) are consistent with the measured variability in the transmission of acoustic signals through the ocean volume. The observations suggest that the statistics of acoustic fluctuations can be used to monitor the statistics of ocean finestructure. Further, fluctuations in the travel time are a measure largely of temperature fluctuations along the acoustic path, and the difference in reciprocal travel times is a measure of the component, along the path, of current and current shear. Such acoustic measurements give integrals along the paths, with the attendant advantages and disadvantages as compared with the traditional 'spot' measurements; precision promises to be comparable or better.

OCEANOGRAPHERS differentiate between marine biology and biological oceanography. The former is concerned with biological processes in marine organisms; the latter places emphasis on biological processes in interaction with physical, chemical and geological processes. Biological measurements can serve as sensitive indicators of the source of water masses, of vertical motion and mixing processes, etc. It is in this sense that we distinguish acoustic oceanography from the intensively active field of marine acoustics.

During and directly following World War II, the interaction between the acoustic and oceanographic communities was close, to the benefit of both. Certain noises in the oceans could be traced to biologic origin, such as snapping shrimp. The deep scattering layer was discovered and readily associated with the known diurnal migration of copepods, towards shallow depths at dusk and downward at daybreak. (Details concerning the organisms among the migrating biomass that provide the

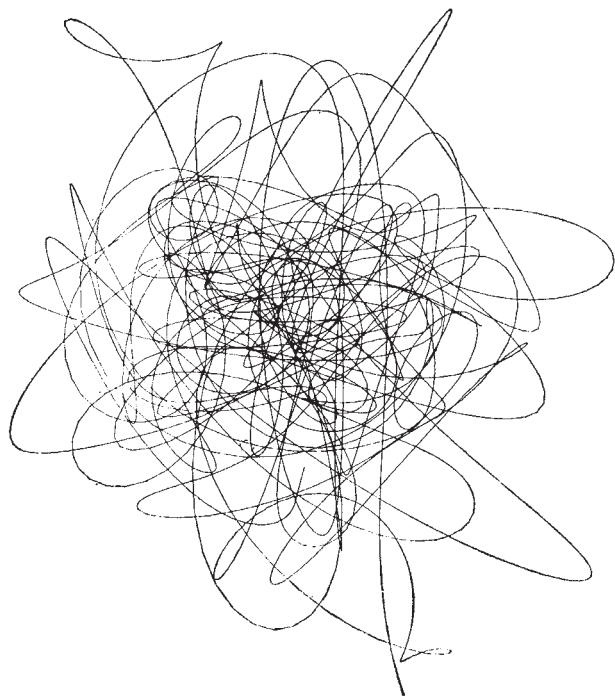


Fig. 1 Phasor diagram of acoustic transmission. The curve consists of straight-line segments through points at 5-min intervals. The distance of a point from the origin (not plotted) is proportional to the intensity, and its bearing angle to the phase of the received signal relative to that at the transmitter. The random walk character of the phasor diagram is associated with interference between many multipaths. The plot was made by computer simulation of the MIMI experimental conditions.

acoustic scattering cross section are complex and not totally solved even today.) The 'afternoon effect' of sharply reduced sonar ranges was explained quantitatively in terms of radiative warming of the near-surface waters. The existence of an acoustic wave guide (the SOFAR channel) was proposed on the basis of the known interior distribution of temperature and pressure, and dramatically demonstrated soon thereafter.

It was then generally accepted (primarily because the oceans are so transparent to sound and opaque to light) that sound propagation would develop into a primary tool for probing and monitoring the oceans. But by and large, the close cooperation between oceanographers and acousticians did not continue. We think this might have to do with the fact that the acoustic community became strongly concerned with the large variability in sound propagation; 20 dB fades are by no means uncommon. And here the oceanographic community was of no help. Oceanographers were preoccupied with very large scale, long time processes, d.c. oceanography so to say, that have little bearing on acoustic variability. What was needed here was a description of the ocean structure on a small scale, and the instruments for making the required measurements had not been developed.

And so the two communities went their own ways, to the detriment of both. The oceanographers were so preoccupied with their wind-spun oceans that they neglected the mesoscale and small-scale processes which contain (as we now know) 99% of the kinetic energy. The acousticians invented their own oceans, borrowing the concept of a homogeneous isotropic turbulence which had served as a successful model for explaining the fading of radio waves in the ionosphere. Temporal fluctuations were interpreted in terms of a frozen ocean structure drifting by (Taylor's hypothesis). But the ocean is neither homogeneous nor isotropic (except possibly on scales less than 1 m. Further, Taylor's hypothesis supposes that the mean currents (winds) exceed the intrinsic velocities of the structure. But in the oceans, the phase velocity of internal waves, for example, is typically 1 m s^{-1} , and the mean motion typically

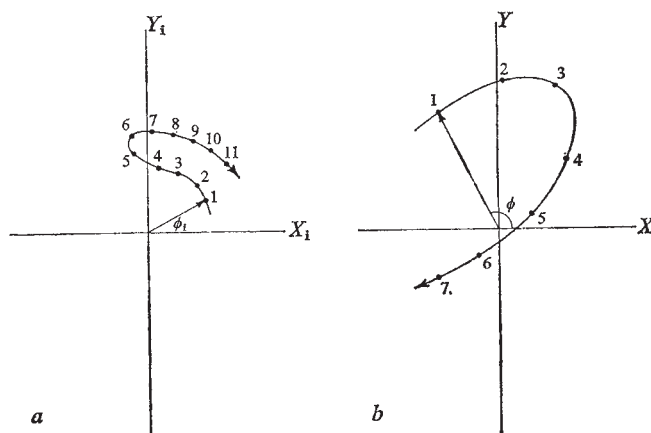
0.1 m s^{-1} . So there is no justification for Taylor's hypothesis, yet it remains ingrained in the acoustic literature (A. W. Ellinthorpe has pointed out to us that most of the existing acoustic observations were carried out from moving vessels, and then the application of Taylor's hypothesis is entirely justified.)

In the 1950s it became apparent that discrepancies between successive measurements at repeatedly occupied oceanographic stations were not always the result of instrumental error. Since then, the development has been rapid. Ocean eddies with time-scales of several months and space scales of 100 km were found to be associated with 100 times the kinetic energy of the general circulation. At frequencies between a few cycles per hour (the Brunt-Väisälä frequency) and a cycle per pendulum day (the inertial frequency) internal waves dominate. Their vertical amplitude is remarkably large, 10 m in the upper oceans, up to 100 m at abyssal depths. The time was ripe to inquire whether internal waves could account for some of the measured variability in the propagation of sound through the ocean.

A long-range sound transmission experiment

In 1966 Steinberg and Birdsall conducted a pioneering sound transmission experiment across the Straits of Florida¹. Interpretation was made difficult by shallow-water effects, and subsequent efforts were shifted to a 1,250 km path from Eleuthera to Bermuda². An essential feature in these experiments (called MIMI for the Miami-Michigan participation) was that they gave continuous observations over many months, and this opened the way for a geophysical interpretation. In essence, the experiment consisted of transmitting a 406-Hz continuous wave signal, and of recording the relative phase and intensity of the received signal using a perfectly synchronised 406-Hz oscillator. The resulting time series of acoustic phase and intensity are dominated by occasional fade-outs and phase jumps which are the result of interference between the many paths from source to receiver (Fig. 1). This is an interesting problem in random-walk statistics, but unfortunately the ocean is involved only in a limited way: the determination of a single parameter, $\langle \phi_i^2 \rangle$, the mean-square rate of phase along any one single path (Fig. 2) (ref. 3). The parameter could in principle be measured directly if such a single path could be isolated from all other paths by a suitable directional antenna. But this is not possible, because each 'deterministic' path is split into 'micro-multipaths' (Fig. 3).

Fig. 2 Phasor diagrams for single path (a) and multipath (b) acoustic transmissions. The Cartesian coordinates designate the in-phase and quadrature components of the received pressure signal, at time 1, 2, ..., with the vector indicating the positions at time 1. ϕ_i and $I_i = 10 \log(x_i^2 + y_i^2)$ are the single path phase and dB intensities, and similar ϕ and I for the multipaths. Generally $\theta(t)$ varies more rapidly than $\phi_i(t)$, and $I(t)$ is more variable than $I_i(t)$. Passage of the vector trajectory near the origin (between times 5 and 6 in the right figure) is associated with a fade-out and a rapid change of phase by almost $\pm 180^\circ$, the sign depending on which side of the origin is passed (minus in the figure).



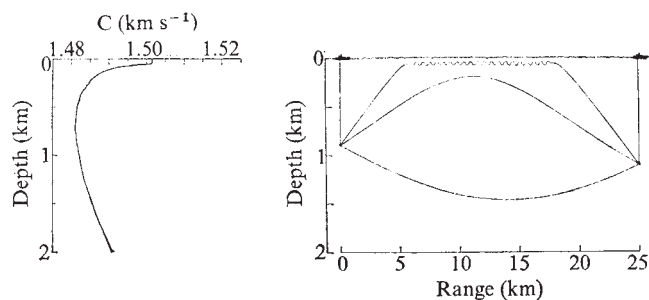


Fig. 4 Sound speed as a function of depth and associated refracted ray paths. A surface duct path is also shown (from ref. 16).

(ϕ) is 1.6 cycles, the value computed from the ocean model is 0.8 cycles. Again there are no loose parameters. But, the measured intensity fluctuations greatly exceed the computed fluctuations, particularly at high frequencies. The discrepancy has not been explained. It could be the result of ocean finestructure over the seamount which differs from typical open ocean conditions, although environmental measurements did not indicate a marked difference.

Acoustic experiment in the Azores

Ellinthorpe's¹⁰⁻¹² ambitious transmission study AFAR in the Azores involved a source at 550 m depth and a receiver at 750 m depth separated by 35 km. The refracted ray is limited by the source depth above, and the sound velocity maximum associated with the Mediterranean outflow at 800 m below. The experiment was conducted over a broad range of acoustic frequencies, and involved an intense program of ocean monitoring.

Interpretation of Ellingthorpe's experiment has gone well beyond the analysis of variance. For this purpose we need to refer to systematics developed by Dashen¹³. According to this systematics the signal statistics fall into five distinct classes, depending on the values of two parameters (Fig. 3). The diffraction parameter $\Lambda = (2\pi)^{-1} R_F^2/L_V^2$ depends on the Fresnel radius R_F and the vertical coherence scale L_V , of the inhomogeneities. The strength parameter $\Phi^2 = \langle [q \int (\delta C/C) ds]^2 \rangle$ depends on the perturbations in sound velocity integrated along the ray path, with q designating acoustic wavenumber. In the lower region of Fig. 3, simple geometric optics and the Rytov extension apply. Ray paths undulate, but they are not split into micro-multipaths. Accordingly, the transmission time of an acoustic pulse will wander, but the pulse shape remains essentially undistorted. In the upper region the deterministic ray path is split into incoherent micro-multipaths, a transmitted pulse is severely distorted and spread, and C.W. intensity fluctuations are saturated in consequence of interference between these paths.

The Azores experiment is in an intermediate region where the intensity is only partially saturated. This, as with other inferences,

is in rather remarkable agreement with observations¹³, and give support to the hypothesis that the fluctuations are the result of internal waves. But most of the first-order statistics do not distinguish sensitively between an internal wave model and the Taylor hypothesis that the inhomogeneities are convected. But, some second-order statistics, such as the intensity covariance, are sensitive to this assumption, and the results are inconsistent with the Taylor hypothesis and consistent with an internal wave model¹³.

An experiment off California

We are preparing for a joint experiment¹⁴ to explore further the transition from unsaturated to saturated forward scattering. We will use multiple moored sources and receivers 250 nm offshore from southern California under a simple ocean geometry, that is, deep water and no seamounts. The moorings will be submerged and self-contained; the penalties for this relatively simple arrangement are the need for Rubidium clocks at each mooring and the need to monitor the mooring motions.

In terms of Dashen's systematics (Fig. 3), the planned diversity of frequencies (220 and 2,250 Hz), ranges (25–125 km), and paths (upper and lower) will cover the transition from geometric optics to coherent micro-multipath (2,250 Hz, SIO) and the transition from geometric optics to incoherent micro-multipath (220 Hz, WHO). Associated ship-based listening stations will extend the 2,250 Hz observations to the incoherent micro-multipath regime. The experiment should provide a crucial test as to whether internal waves impose the fundamental coherence limits (space and time) to marine acoustics.

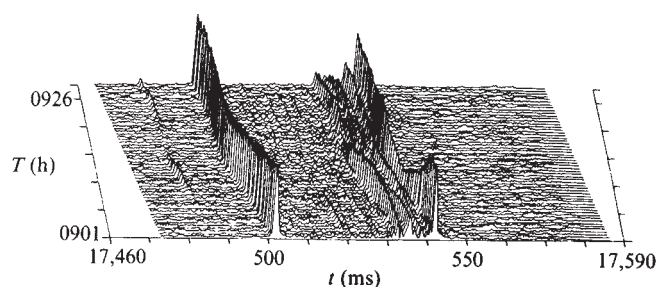
A second (and related) goal of this work is to assess the role of acoustic oceanography in synoptic monitoring of the oceans on a mesoscale (~100 km). It has become increasingly apparent in the last several years that most of the kinetic energy of the oceans is contained in mesoscale processes¹⁵. Mesoscale 'eddies' are analogous to atmospheric storm systems (~1,000 km scale). This disparity in spatial scales makes the preparation of an 'ocean weather map' more difficult than preparation of the corresponding atmospheric map. A reverse disparity in time scales of ocean and atmospheric weather (~60 d compared with ~3 d) does not offset this difficulty. Acoustic oceanographic methods offer one possible mean for gathering data needed to map ocean weather.

The 2,250-Hz work off California is a preliminary step towards acoustic ocean monitoring. We will take advantage of the fact that the mean sound velocity profile gives one direct upper ray path and one direct lower ray path between two points located at about 1,000 m depth 25 km apart (Fig. 4). Travel time for reflected paths are well separated in time from those for direct paths and can easily be discriminated against. Very encouraging results from ship-based transmissions using this geometry (Fig. 5) have been obtained by Worcester¹⁶.

Changes in the oceanographic variables—temperature, salinity, and current—contribute to changes in travel time. Simultaneous transmission of pulses in opposite directions differentiates between current effects and those of temperature and salinity. Worcester has attained an absolute travel time precision of 10^{-4} s and a relative precision of 2×10^{-6} , using modified matched filtering techniques. A fluctuation in travel time by 10^{-4} s corresponds to a temperature fluctuation of 10^{-2} °C, a salinity fluctuation of 10^{-2} ‰, or a current of 1 cm s^{-1} integrated along the ray path. The 25 km averaging accomplished by the adopted mooring spacing is appropriate for detecting structures of eddy dimensions. We would, therefore, expect a distinctive travel time signature from the drift of a mesoscale eddy past the moorings. Identification of the contribution of individual variables to the overall travel time effect is as yet unsolved. A complete solution of this inverse problem probably requires data from many ray paths, each giving the result of its own vertically and horizontally weighted mean.

An environmental measurement program will accompany the acoustic work to compare results of the acoustic measurements

Fig. 5 Arrival amplitudes at 30-s intervals after high resolution processing (from ref. 16). Time is relative to the start of the transmitted pulse. Note the internal wave-caused relative complexity of the later, upper path arrival.



with data from traditional oceanographic instruments. A measure of the degree to which acoustic monitoring fills an unoccupied niche is the difficulty one faces in using conventional instruments to 'check' acoustically derived data. Even for our small, two-mooring experiment, we have not found it practical to monitor both acoustic paths with moored instruments. We will instead use conventional current and temperature meters to make continuous measurements along the simpler lower ray path only. Frequent upper path measurements will be made with instruments lowered from ships. Comparison of acoustic and 'spot' measurements will be limited largely to subinertial and tidal frequencies.

Acoustic monitoring of ocean weather has an advantage over conventional methods: by measuring between stations rather than at stations the information collected goes up with station number like $n(n-1)$ rather than n (allowing for reciprocals). There are no major technological problems. Indeed, for these subinertial frequencies one would dispense with the bothersome mooring position-keeping. Surface floats could allow for radio

control of clocks. The automated analysis of acoustic signals is a formidable task, but existing microprocessor technology is adequate. Experiments over the next few years will go far towards determining the feasibility of 'listening' to undersea weather.

- ¹ Steinberg, J. C. & Birdsall, T. G. *J. acoust. Soc. Am.* **39**, 301-315 (1966).
- ² Clark, J. G. & Kronengold, M. *J. acoust. Soc. Am.* **56**, 1071-1083 (1974).
- ³ Dyson, F., Munk, W. H. & Zetler, B. *J. acoust. Soc. Am.* **59**, 1121-1133 (1976).
- ⁴ Garrett, C. J. R. & Munk, W. H. *J. geophys. Res.* **80**, 291-297 (1975).
- ⁵ Cairns, J. L. & Williams, G. O. *J. geophys. Res.* **81**, 1943-1950 (1976).
- ⁶ Wunsch, C. *J. phys. Oceanogr.* **6**, 471-485 (1976).
- ⁷ Munk, W. H. & Zachariasen, F. *J. acoust. Soc. Am.* **59**, 818-838 (1976).
- ⁸ Ewart, T. E. *J. acoust. Soc. Am.* **60**, 46-59 (1976).
- ⁹ Desaubies, Y. J. F. *J. acoust. Soc. Am.* **60**, 795-803 (1976).
- ¹⁰ Ellinthorpe, A. W. *NUSC Tech. Doc.* 4551 (1973).
- ¹¹ Ellinthorpe, A. W. & Freese, H. A. *NUSC Conf. Tech. Rep.* 4647 (1973).
- ¹² Ellinthorpe, A. W. & Krulisch, A. H. *NUSC Tech. Memo. No.* TE-105-75 (1975).
- ¹³ Flatté, S. (ed.), Dashen, R., Munk, W., Watson, K. & Zachariasen, F. *Sound Through a Fluctuating Ocean* (in the press).
- ¹⁴ Principal participants are W. H. Munk, G. O. Williams, P. Worcester, B. Zetler (Scripps Inst.), R. Porter, R. Spindel (Woods Hole Oceanogr. Inst.) & F. Spiess (Scripps Inst.).
- ¹⁵ MacLeish, W. H. & Price, J. (eds) *Oceanus* **19**, (WHOI, Woods Hole, spring 1976).
- ¹⁶ Worcester, P. F. submitted to *J. acoust. Soc. Am.*

The life of the open sea

John D. Isaacs

Scripps Institution of Oceanography, Institute of Marine Resources, University of California, La Jolla, California 92093

"Third Fisherman—Master, I marvel how the fishes live in the sea.

First Fisherman—Why, as men do a-land; the great ones eat the little ones."

Pericles

THE living creatures of the near-surface waters of the open sea survive in one of the highly demanding and unforgiving habitats of the planet. Yet here a great and varied assemblage of creatures has evolved in or adapted into a vast realm that offers neither refuge nor obvious environmental niches to its myriad of inhabitants. The student of this region is strongly challenged by the very subtlety and obscurity of the adaptations and modes of existence, and often is astonished by his findings.

I will briefly review here some major aspects in the present understanding of the open sea, the nature of the forces and conditions that have set the range and compass of its progeny and will attempt to point out the curious and poorly understood steps that allow a relatively large population of highly developed, highly active animals to ensue from a food web initiated by microscopic plant cells, thinly dispersed in a dense fluid medium. I will speculate on defects in our present approaches towards understanding marine trophic (food web) systems and on an example of overlooked possibilities. Within the discussion, I will touch only briefly on adjacent interrelationships, coastal waters, shallow seas, coral shoals, and the environments of the deeper bathypelagic and benthic life. This short shrift to the boundaries of the environment that I discuss should be taken as a measure of the limitations of space rather than of importance, for the children of Tethys and the flocks of Nereus are, of course, no more isolated from the conditions and inhabitants of other realms than the disciplines of modern marine science can be cleaved one from another.

A properly structured discussion of any living system will begin with the processes by which the inorganic chemical and energetic sources enter and recycle in the system; and proceed with the living autotrophic entities that catalyse the elaboration of initial organic material from these lifeless sources. It will then describe the myriad and often

bizarre living forms and trace the Byzantine paths by which this organic material is utilised for its substance, processed for its energy, and finally returned again to the inorganic reservoir. I point out some developments that modify the classical visualisation of these steps.

Modern research largely affirms the broad aspects of the conventional concepts of the initial input of inorganic nutrients into the near-surface waters of the ocean. In this picture the special chemical requirements of phytoplankton as well as those common to almost all plant life are abundantly supplied only in those parts of the sea where some dynamic process provides the mechanical work required to recycle nutrient-rich water from deeper levels where they have been released, into the lighted upper layers where they are used. But, these sunny well-ploughed pastures of the sea, where luminous and chemical needs coincide, account for only a few per cent of the ocean's primary productivity, for they are of limited area. They exist mainly along eastern boundary currents, in the Antarctic, and at the equator, where the processes of upwelling hold sway. Because of the vast extent it is the great oceanic areas that account for most of the total productivity some 80%, despite the low production rates in these leaner waters¹.

It is clear, however, that the classical picture of the primary trophic inorganic steps needs to be modified in detail. Certain algal colonies (blue-greens) elaborate significant quantities of the essential nutrient, inorganic chemically-bound nitrogen². The recycling of nutrients within shallow seas and shelves, in coral atolls and other environments apparently contributes more to their productivity than previously supposed. In such environments, there may be only very low levels of plant nutrients dissolved in the water, for here recycling and uptake proceeds apace³, as in jungle soil. Recycling may also be internal, within animal tissues that bear captive plant symbionts, as in the case of corals and the *Tridacna* (giant) clams⁴. Hence classical chemical measurements of dissolved nutrient levels and inferences of productivity may be wholly misleading. Another process that modifies the classical portrayal seems to take place in the very clear waters of central oceanic gyres. Here in the uppermost

levels of the thermocline is a layer dense in chlorophyll, although the light level of less than 1% of surface illumination is below the minimum previously assumed necessary for net phytoplankton growth⁵. Apparently excess chlorophyll along with ancillary pigments or other adaptations allow phytoplankters to grow at the poorly illuminated diffusing upper edge of the nutrient-rich submerged waters—a process probably only general in the relatively steady conditions and clear near-surface waters of central oceanic regions.

Thus in the light of modern research, the general elements of the classical portrayal of the first step of the trophic system, primary productivity, survive quite well. Subsequent aspects of the conventional food-web picture seem to require more modification, however.

For example, dissolved organic material (DOM, the 'gelbstoff', refractory highly oxydised organic substances of high molecular weight) may assume an increasingly important trophic role. This material is undoubtedly the greatest of the organic reservoirs within the ocean, 200 times the total phytoplankton and particulate organic material⁶. Furthermore, its average age (3,400 yr) (ref. 7) argues that it is only utilised or removed slowly and perhaps only in some special conditions. Slow recycling of this great reservoir may be conducted by certain recently isolated bacteria of deep water which can utilise it⁸. It may become rapidly available along the hydrothermal areas of deep rift zones at oceanic ridges, where heat, pressure, and brief reducing conditions may transform it into more available chemical forms. Wherever it is that DOM is recycled, food input could be very large. For example, the rate at which DOM becomes available in the North Pacific could exceed the primary productivity there by a factor of 10 (ref. 9).

Other modifications of the trophic picture emerge apace, and I will sketch only a few. Remarkable discoveries of the nature of metabolites and organic pathways include the demonstration of a bewildering range of halogenated organic products (a chemical province recently thought to be the chemists' exclusive domain) elaborated by marine plants for various defensive, offensive, and possibly communicative purposes¹⁰; the dominant role of waxes in major parts of the food webs of the open sea and coral reefs¹¹; the role of the ubiquitous behaviour of the active pelagic zooplankton and small nectonic (more strongly swimming) organisms in performing a diurnal round-trip migration from the surface to mid-depth, thus providing a number of functions—including leading them to productive areas¹²; the demands of energetics that force large active predators preferentially to select surprisingly small organisms for their major food resource¹³; the well-nigh invisible environmental niches provided for apparently fully equivalent species by their different responses—not to conditions—but to the time rate of change of conditions¹⁴; the remarkable chemotactic-motor trail-following behaviour of some marine organisms¹⁵; and many others.

In this last example, sergestid shrimp have been shown to follow a fine thread-like trail of amino acids with ultimate swimming speed and preciseness, an astonishing display of chemotactic motor control.

These examples epitomise the surprising discoveries that the observation of the forms, behaviour, and processes of life of the sea will ever yield up to its students. At the same time the same examples may also indirectly point to one of the deficiencies of the underlying observational, scientific, and philosophical approach to the trophic pathways of the open sea and perhaps to other aspects of marine biology as well.

The careful and thoughtful observations and measurements of living forms, behaviour and pathways have taught much important understanding in forest and grassland¹⁶ and in some estuarine, near-shore and other

environments. Yet in the open sea, and despite the inherent fascination of naturalists towards the activities of the creatures of the sea, many of the dominant activities and pathways are fundamentally not discernible. Major pathways of detritus, reproductive products, growth and dissolved organic material are virtually invisible, yet the delusion commonly persists that with only sufficient collections, observations, measurements and analysis, the full dynamic nature of the food webs of the open sea eventually will stand revealed. I will enlarge briefly on this, for it may well be the principal diastema in our philosophical approach to the trophic systems that dominate two-thirds of this planet.

The flow of food from initial microscopic levels of the open sea to higher forms such as fishes, squid, and marine mammals seems to be dominated by quite a different process from similar fluxes on land or near-shore and shallow aquatic environments, where the principal inputs are by large plants¹⁷. In the pelagic realm a sort of food web composed of the progeny of many species seems to constitute a major pathway of material to the larger forms. This pathway is dominated by growing larvae and juveniles, which suffer immense mortality in their role. It is a sort of huge polyspecific extension and exacerbation of the 'Faginism' observed in some cases. For example, in lakes that have been stocked with one species of predacious fish, a food web develops from successive predation of juveniles by larger juveniles ultimately supporting a few reproducing adults. Such a system is in sharp contrast with other food webs where large plants provide the initial input, large herbivores are supported, the populations are dominated by adults, and mortality is low.

Undoubtedly this system has developed in the open sea for access of the larvae and juveniles to the abundant microscopic and near microscopic food at the commencement of the food web. It is responsible for the immense flow of food material as reproductive products (egg, sperm, and larvae) from the higher forms down towards the beginnings of the system, and may be largely responsible for the very existence of higher forms in the open sea.

It is interesting to speculate on the role of those marine creatures that take from but do not contribute to these essential pathways. The terrestrial expatriates (marine mammals, birds and reptiles) and some sharks which bear large young, depend entirely on the peculiar pathways of the pelagic food web but escape from both the associated contribution to and the predation within the system upon which they are dependent. Any substantial dominance of these forms may thus lead to collapse of the entire higher pelagic system. This fundamental instability of the higher levels of the pelagic food web may illuminate events in the palaeontological record where explosive proliferation of highly-developed (and apparently viviparous) pelagic predators culminated in their own universal mass extinctions and that of many other higher pelagic creatures¹⁸.

The present marine mammals and tuna-porpoise policies need to be restudied in light of this possibly fundamental long term ecological instability of the pelagic system.

I will conclude with an added example of each of two principal themes on our understanding the conduct of living system of the open sea, one philosophical and one observational.

First, there are opportunities to view trophic systems by overleaping the details of processes with the same philosophy that enables us to coalesce a microcosm of molecular kinetic complexity into a single meaningful measure—temperature. Such an approach complements the more detailed naturalistic descriptions and measurements. One such model views organisms as operators that continually distribute and redistribute resources into a few specific reservoirs by appropriate transfer coefficients¹⁹. These reservoirs are, most simply, living material, dead

retrievable material, dead irretrievable matter, and reproductive products. The approach generates useful and meaningful expressions of potential relative biomass of trophic types and other trophic relationships, and seems to explain a number of anomalies, such as the frequent observation that the biomass of predacious forms exceeds that of the herbivorous forms with which they are in an apparent equilibrium. It also provides reasons for both the general immunity of marine organisms from the bioaccumulation of trace substances in the food web (including pollutants); and the vulnerability of marine birds and mammals to the bioaccumulative process.

As a final speculative example, aggregation of the active creatures of the open sea, such as schooling, is an almost universal feeding, defensive, predatory, or mating strategy. Indeed, the existence of both baleen whales and marine fisheries is almost wholly dependent on species that display strong schooling behaviour. Nowhere is schooling more conspicuous than in the small filter feeding fishes, the clupeoid sardines, anchovies, herrings and their relatives. It has long been known that schools of such fish are strong scatterers of underwater sound and easily detected by sonar devices. The strong sonar reflection of these fish is due to the abrupt discontinuity in sound velocity at the boundaries of their gas-filled swim bladders. Although underwater sound of the frequency and wave length in 'fish-finders' responds to the presence of a swim bladder as though it were a discontinuity in the water, lower frequency sound will respond to the swim bladder as a property of the medium. This is quite properly analogous to the scattering of light by fog and the penetration of such obscuration by radar.

Thus, 1,000-cycle (1 kHz) underwater sound with a wave length of almost 2 m will envelope a number of fish in the school per wave length. Simple calculations show that the velocity of low-frequency sound within a fish school can be as low as one-third the velocity in sea water. Such a change in sound velocity is orders of magnitude greater than that engendered by variations in the ordinary physical-chemical conditions of temperature, salinity, and

pressure. Thus to low frequency sound the fish school is a wholly unique object, and such sound must focus abruptly within the schools. A consequence of the schooling of small fish is thus the creation of an immensely powerful, complex and continuously pliant lens for the focusing of low frequency sound, a lens that is composed entirely of detectors of such energy! Here is an example of the development of certain qualities (for example, a hydrostatic device) and particular behaviour (for example, schooling) under quite disparate requirements interacting to form a wholly new advantageous property, which could not be approached at the level of mutation of the individual organism. In this case the serendipity is the creation of a unique living lens to magnify and respond to the sounds of distant predators!

In this brief resume of the life of the open sea, I have suggested some of the modifications of the classical picture emerging from the observations and the burgeoning new insights into its forms and processes; and some of the major deficiencies in the present approaches and philosophy.

I have also speculated on examples of the possible future advancements of understanding that the instructress, Nature, may ultimately teach us within the reaches of the great realm that encompasses two-thirds of the surface of this planet, the open sea.

- 1 Ryther, J. H. *Science* 166, 72 (1969).
- 2 Goering, J. J., Dugdale, R. C. & Menzel, D. W. *Limnol. Oceanogr.* 11, 614 (1966).
- 3 Fogg, G. E. *Algal Cultures and Phytoplankton Ecology* 82 (University of Wisconsin Press, Madison, Milwaukee, and London, 1965).
- 4 Johannes, R. E. *et al. Bioscience* 22, 541 (1972).
- 5 Venrick, E. L., McGowan, J. A. & Mantyla, A. *Fish. Bull.* 71, 41 (1973).
- 6 Strickland, J. D. M. *A. Rev. Microbiol.* 19, 127 (1965).
- 7 Williams, P. M. *Nature* 224, 256 (1969).
- 8 Williams, P. M. & Carlucci, A. F. *Nature* 262, 810 (1976).
- 9 McGowan, J. A. & Williams, P. M. *J. exp. mar. Biol. Geol.* 12, 187 (1973).
- 10 Fenical, W. H. *J. Phycol.* 2, 245 (1975).
- 11 Benson, A. A. & Muscatine, L. *Limnol. Oceanogr.* 19, 810 (1974).
- 12 Isaacs, J. D., Tont, S. A. & Wick, G. L. *Deep-Sea Res.* 21, 651 (1974).
- 13 Isaacs, J. D. & Wick, G. L. *Spl. Publ. Mar. Biol. Ass. India*, p. 193 (1973).
- 14 Silver, M. W. thesis, Univ. California (San Diego, 1971).
- 15 Hammer, P. & Hammer, W. M. *Science* 195, 886 (1977).
- 16 Woodwell, G. M. *Scient. Am.* 223, 64 (1970).
- 17 Isaacs, J. D. *Bull. Sth Calif. Acad. Sci.* 75, 220 (1976).
- 18 Fischer, A. G. & Arthur, M. A. *Spec. Publ. No. 25 Soc. Econ. Pal. Min.* (in the press).
- 19 Isaacs, J. D. *Mar. Biol.* 22, 97 (1973).

Diversity and faunal composition of the deep-sea benthos

Torben Wolff

Zoological Museum, University of Copenhagen, Denmark

Improved techniques have demonstrated a remarkably high species diversity in the deep sea. Long-term environmental stability has resulted in establishment of diverse communities whose species have become biologically accommodated to each other. The deep-sea meiofauna (about 60–500 µm) mainly consists of foraminiferans and nematodes. The macrofauna is dominated by polychaetes, peracarid crustaceans, and bivalves. The major components of the megafauna are brittle stars bathyally, large scavenging fish and amphipods abyssally, and amphipods hadally. Biomass and numerical density decrease rapidly with distance from land. Biomass seems to be of the same order of magnitude for the three size groups, while meiofaunal density in the upper bathyal zone may be up to 1,400 individuals per 10 cm² and macrofaunal density 20,000 individuals per m².

In the last decade, modern techniques have led to many remarkable discoveries regarding life and life conditions in the deep sea. Rather than attempt a general brief survey of new knowledge gained in almost all fields of deep-sea biology, I have preferred to concentrate on a limited number of topics which represent some of the most striking advances in our conception of deep-sea life.

Species diversity

Until the 1960s it was the general experience that with increasing depth and distance from land the benthic fauna consisted of fewer individuals belonging to fewer species. The recent introduction in deep-sea research of the anchor dredge¹, the epibenthic sled², and careful washing of the samples through fine-meshed sieves drastically changed this concept. Greatly increased numbers of individuals now became available, numerically increasing the number of species, and also allowing the taxonomists to recognise that the samples contained numerous closely related species.

The high species diversity was first demonstrated in the

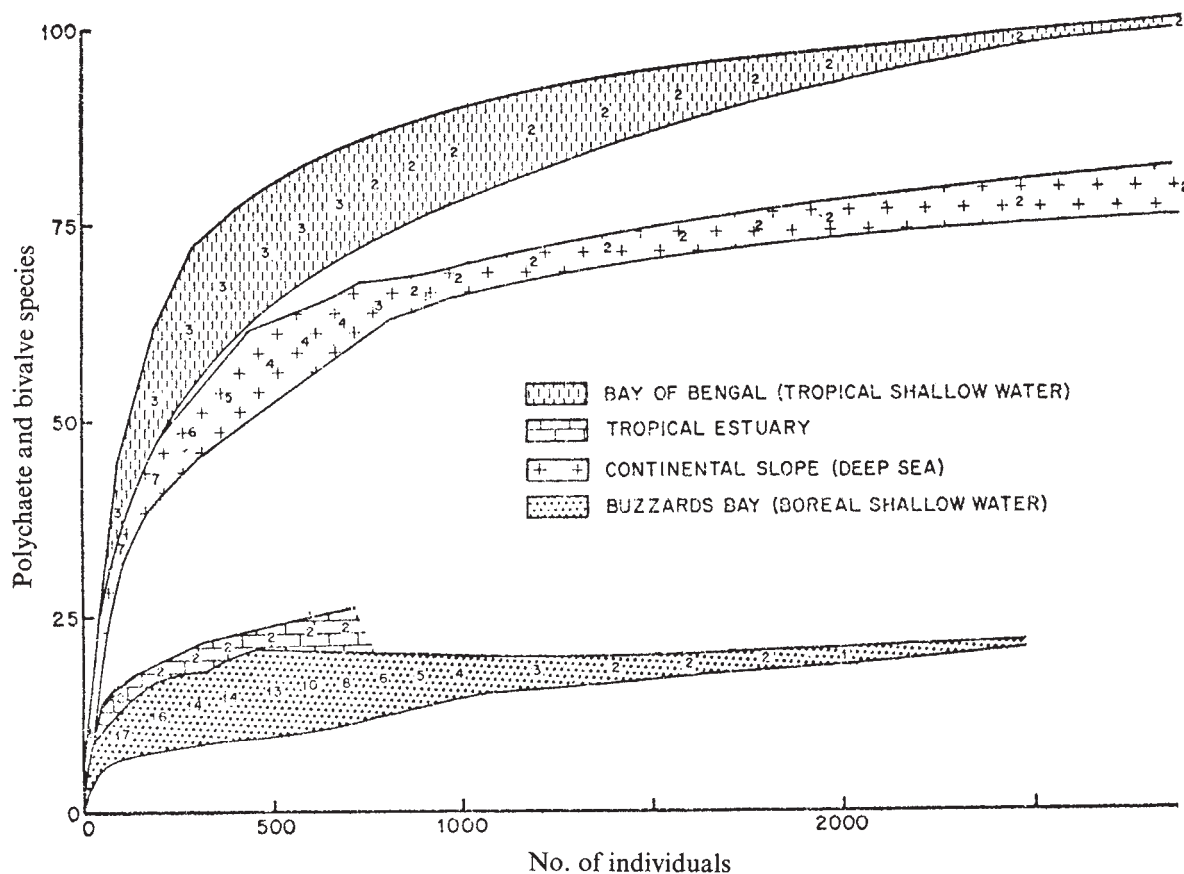


Fig. 1 Range for diversity values found for three shallow-water and two combined deep-sea environments (the latter in the NW and the SW Atlantic, 300–2,500 m). The figures in the bands indicate number of samples. (Slightly modified from ref. 6.)

north-western Atlantic². Samples subsequently taken off western Europe, in four basins off West Africa and in the Brazil Basin, collected and processed with similar methods, show the same general pattern.

For instance, three sled samples from the north-western Atlantic from 1,400, 2,900 and 4,700 m contained 25,200, 12,100 and 3,700 individuals belonging to 365, 310 and 196 species, respectively². The hitherto most successful deep-sea sample was taken with a large trawl in the East Pacific at 3,600 m and contained 2,100 individuals belonging to 132 species³. Among the dominant groups, the greatest diversity is invariably found in peracarid crustaceans, followed by polychaetes and bivalves, while ophiuroids are always represented by few species. If there is one species of brittle stars per 1,000 individuals, there are four species of bivalves, 10 of polychaetes, and 17 of peracarid crustaceans⁴. Also harpacticoid copepods, an important component of the deep-sea meiobenthos, show high diversity⁵.

In order to compare deep-sea diversity with that in shallow water, soft-bottom samples were obtained from, among other places, a boreal and a tropical estuary and a boreal and a tropical marine environment, and processed with the same techniques. Based on the polychaete-bivalve fraction of the fauna, which accounts for about 80% of the total fauna in most infaunal samples, and using a rarefaction methodology which allows direct comparison of samples with differing numbers of specimens, Sanders^{6,7} showed that only in the physically stable, shallow, tropical marine environment was there a diversity as high as that of the deep sea (Fig. 1). Conversely, a low diversity is characteristic in environments where physical parameters (for example temperature or salinity) fluctuate greatly.

Basic to Sanders' stability-time hypothesis is the idea that long-term environmental stability has allowed a highly refined competitive niche differentiation. Since food is the

most important limiting factor in the deep sea, each species may have evolved a high food specialisation, thereby avoiding competitive overlap with other species of the community.

Although they do not deny the presence of niche specialisation, Dayton and Hessler⁸ believe that continuous and intense biological disturbance by predators is the main element responsible for controlling population size and maintaining high diversity. They find that in the deep sea a high degree of food-niche specialisation is unlikely. There can be no sharp demarcation between deposit feeders and predators, because in order to survive, both categories must possess a certain flexibility in their food choice. An appropriate term for animals of such a broad feeding-type is 'croppers', that is animals of any size which consume live food, either alone or in combination with dead or inorganic materials. The pressure exerted by croppers on populations of smaller animals reduces the probability of competitive exclusion of the latter and allows a high overlap in the utilisation of food resources.

Grassle and Sanders⁹ think that in view of the known life history characteristics of deep-sea species, predation cannot be the principal mechanism for controlling population size. Recent investigations of numerous individuals belonging to single species of the tan-aid family Neotanaisidae, of the isopod family Desmosomatidae, and of protobranch bivalves showed small brood sizes and that each population is characterised by a much higher proportion of adults compared to young stages than encountered in comparable shallow-water species. Moreover, growth rate is slow and longevity high, as dramatically demonstrated by Turekian *et al.*¹⁰, who by using ²²⁸Ra chronology determined the age of a protobranch deep-sea bivalve, measuring up to 8.4 mm, to be about 100 yr. Small population size, low reproduction and mortality

Table 1 Composition of the macrofauna (% of the total number of individuals) at bathyal depths off Southern California and in the Arctic Ocean, at abyssal depths in the Atlantic and the Pacific, and at hadal depths in the Kurile-Kamchatka, Japan, Kermadec, and Java Trenches.

Gear	Core samplers		Anchor dredge			Epibenthic sled			Trawl
Source	Ref. 13	Ref. 18	Ref. 11			Ref. 2			Ref. 19
Area	Californian basins	Arctic Ocean	North-western Atlantic		Central N. Pacific	North-western Atlantic			Trenches
Depth (m)	1,130–1,230	1,060–2,530	300–3,750	4,400–5,000	5,600	1,400	2,500–2,900	4,700	6,200–10,000
Group									
Polychaeta	76	50	70	55	55	43	5	8	7
Crustacea Peracarida	14	9	7	33	24	17	20	32	5
Tanaidacea	4	3	2	19	18	4	1	1	<1
Isopoda	3	5	1	12	6	6	6	18	4
Amphipoda	5	...	4	2	...	4	3	5	<1
Bivalvia	2	9	13	4	7	18	5	47	19
Echinodermata	3	<1	<1	1	1	11	67	2	57
Ophiuroidea	2	<1	<1	<1	<1	5	64	2	2
Holothurioidea	<1	<1	<1	...	<1	5	3	...	54
Others	4	31*	9†	8‡	12‡	2	3	11§	11
No. of individuals	5,164	167	28,277	681	287	25,242	31,405	3,737	21,589

*Mainly tetraxonid sponges.

†Mainly sipunculids.

‡Mainly oligochaetes, scaphopods, and bryozoans.

§Mainly gastropods.

||Mainly pogonophorans due to a mass occurrence in one sample from the Kurile-Kamchatka Trench.

rates, and proportionately modest numbers in the smaller size groups would neither be expected nor have high survival value in predator-controlled communities or in any environment where transient diversity is important. These characteristics would, however, be expedient where long-term or evolutionary diversity is a product of past biological interactions in a physically stable environment such as that of the deep sea.

In order to arrive at a better understanding of the causes of deep-sea diversity, other methods must be applied, involving, for example, detailed analyses of microdistribution, nutrition, biochemical specialisation, and predator-prey relationships.

Construction of the box core sampler covering an area of 0.25 m² and partitioned into 25 subcores, each 100 cm², has made it possible not only to obtain quantitative data comparable with those obtained by bottom grabs (see below), but also to evaluate the microdistribution.

A comparison of core samples from the oligotrophic sea floor at 5,600 m in the central North Pacific Gyre showed a diversity of polychaetes and crustaceans exceeding that of sled samples from the oligotrophic north-western Atlantic¹¹. This high diversity in an extremely oligotrophic environment indicates that sparseness of food does not affect diversity.

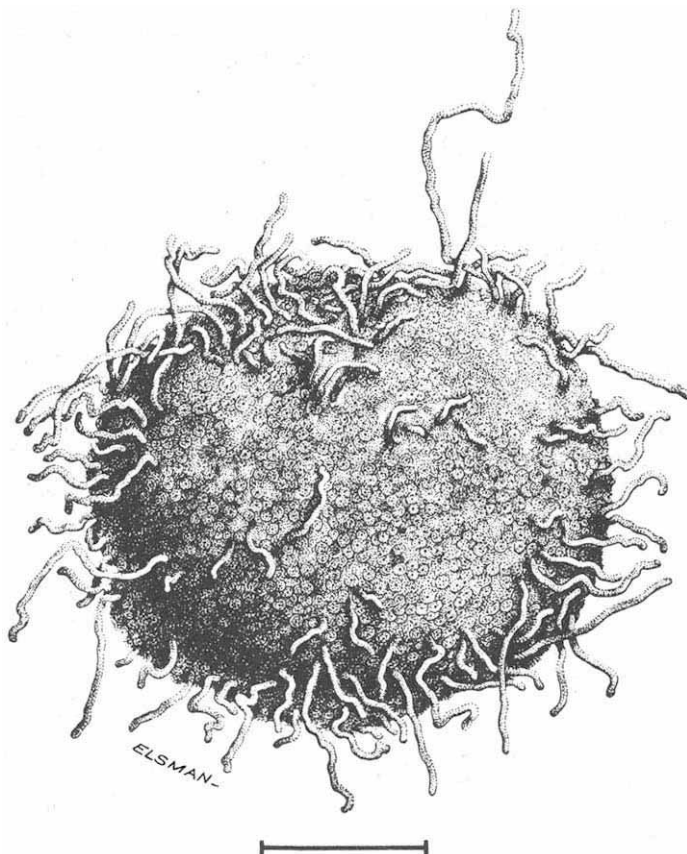
Jumars¹² studied the polychaete fauna in partitioned box core samples from 1,230 m off Southern California and found that a uniform dispersal dominated within cores. Thus, the extremely high diversity observed cannot primarily be due to patchiness or spatial heterogeneity. But, species assumed to be sedentary (small-ambit species) were more diverse than species assumed to be mobile and thus having larger ambits. This indicates that microhabitat specialisation, acting on spatial scales smaller than 100 cm² (the area of the subcores), here has a major role in maintaining high polychaete species diversity. Similar results were obtained by comparing the entire fauna in box cores from two bathyal basins off Southern California¹³.

Composition and biomass of the deep-sea benthos

The grouping of benthic organisms into size classes reflects the sampling and processing techniques. Thus, it is generally accepted that the meiofauna comprises animals which are

retained on screens of about 60 µm and pass through screens coarser than 300–500 µm. The macrofauna is retained on the latter screens, while the megafauna consists of larger animals which cannot be collected adequately with a grab or core sampler for quantitative evaluation.

Fig. 2 One of many komoki from 6,100 m in the central North Pacific. It belongs to a new superfamily of agglutinating foraminiferans which have been found to be extremely abundant in the deep sea. Scale bar 1 mm. (After ref. 21.)



Most of these animals are large enough to appear in survey photographs.

The meiofauna: The number of living foraminiferans varies considerably regionally and with depth. Highest numbers (up to 1,400 per 10 cm² and 1 cm sediment depth) occur at bathyal depths off Southern California, while the mean numbers in the East Atlantic are an order of magnitude lower¹⁴. This is also the case in the Kurile-Kamchatka area, with highest numbers at bathyal and hadal depths and lowest numbers abyssally¹⁵. In contrast to other meiofaunal components (except to some extent the nematodes), a significant proportion of foraminiferans (up to 60%) lives in sediments deeper than the uppermost 1 cm (ref. 14). Calcareous foraminiferans dominate above about 2,000 m, while agglutinating forms are the sole components deeper than 3,500–4,500 m.

Comparisons of number of foraminiferans with the total meiofauna range from 2–27% foraminiferans off Mauritania at 250–4,500 m (ref. 14) to 2–90% (mean 61%) at 250–2,500 m (ref. 16) and 65% at 4,000 m (ref. 17) off North Carolina. In biomass, foraminiferans constitute 15–86% (mean 53%) in the Arctic Ocean¹⁸.

Except for local abundance of foraminiferans, the overall dominating group is the nematodes, in numbers as well as in biomass. Excluding foraminiferans, nematodes constitute 39–96% (mean 85%) of the meiofauna in the 165 compared samples and subsamples so far investigated from 11 areas at bathyal and abyssal depths. Dominance of nematodes does not seem to vary with increasing depth or distance from land.

Second in importance numerically are harpacticoid copepods, followed by polychaetes, ostracods, and crustacean nauplii. Rotifers, oligochaetes, bivalves and tardigrades may be present.

The density of organisms varies greatly, from 0.3 individuals per 10 cm² in the abyssal oligotrophic central North Pacific to 1,000–1,400 individuals per 10 cm² on the East and West Atlantic continental slope. The lowest biomass recorded is 0.1–0.05 g m⁻² surface and 4 cm sediment depth in the abyssal Mediterranean and the abyssal central North

Pacific, the highest is 3.5–6 g m⁻² at bathyal depths in the North Atlantic. At abyssal depths the biomass may be as high as 2.8 g m⁻² (the Norwegian Sea) and is usually 0.2–0.3 g m⁻².

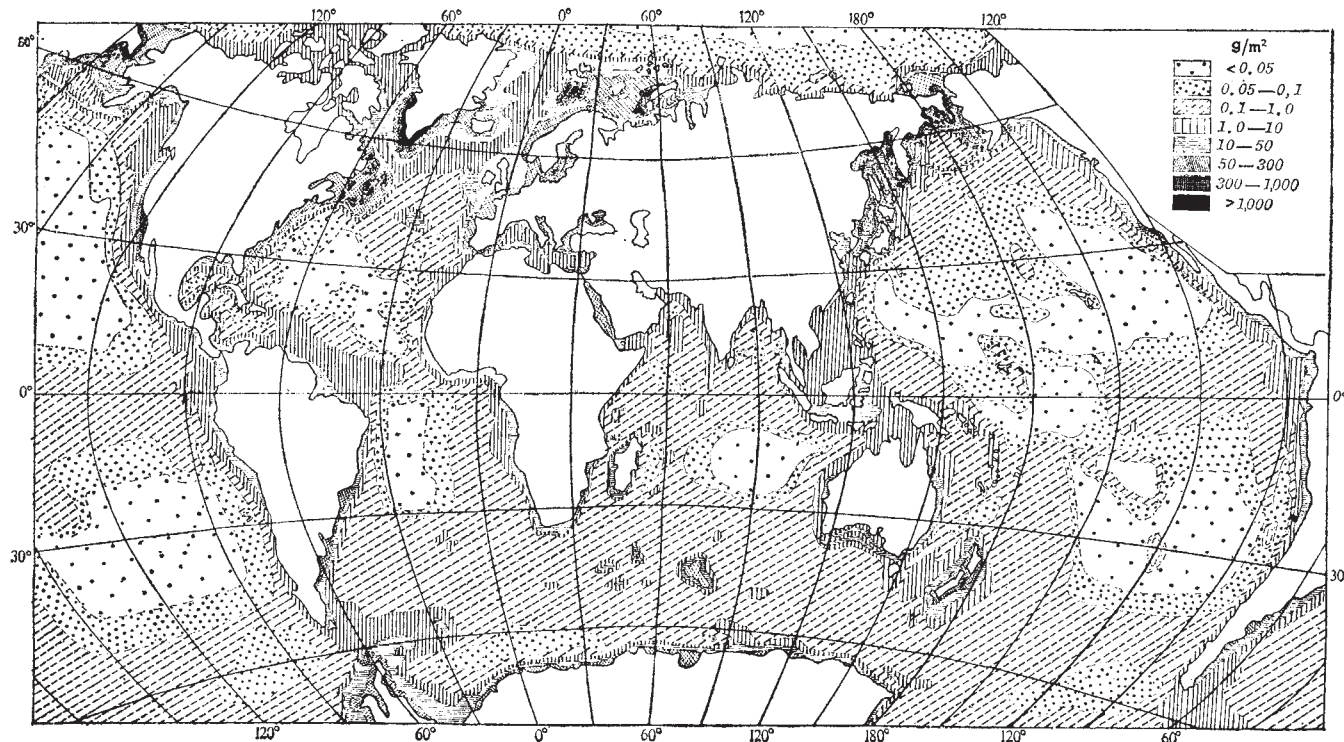
Both numerically and in biomass there is a significant difference between the macro- and meiofauna with increasing depth. The marked macrofaunal decrease is not typical for the meiofauna. As mentioned above, meiofaunal biomass may be remarkably high at abyssal depths. Transect investigations in seven areas of the Atlantic¹⁴ show a maximum decrease of meiofauna biomass of 16–83% mean (50%) between upper slope and 3,000–5,000 m.

The macrofauna: A detailed analysis of 10 box core samples from two basins off Southern California (1,200 m) showed that polychaetes constitute more than 75% of the total number of individuals (Table 1). Next in importance are the peracarid crustaceans, with an almost equal share between amphipods, tanaids and isopods. Rather insignificant are bivalves and echinoderms (mainly consisting of brittle stars). In 75 quantitative samples from mainly bathyal depths in the Arctic Ocean half of all individuals (minus nematodes) were polychaetes (Table 1), followed by tetraxonid sponges (13%), crustaceans, and bivalves.

Also in abyssal communities, polychaetes, crustaceans, and bivalves generally are the main components of quantitative samples. In Table 1 the relative number of individuals in the marginal and oceanic north-western Atlantic and in the central North Pacific Gyre are compared. Polychaetes dominate, although in the oceanic environments they decrease somewhat in importance. The same applies to bivalves, while crustaceans, particularly tanaids, become proportionately more abundant. Combined, the three groups constitute 86–92% of the total fauna.

The above comparisons are based on box core and anchor dredge samples of the infauna. In epibenthic sled samples from the north-western Atlantic, which consist of both epifauna and infauna, a rather similar relative abundance prevails bathyally (Table 1), but at abyssal depths the composition is different. Polychaetes have an

Fig. 3 Quantitative distribution of zoobenthos (wet weight biomass) in the world oceans. (Slightly modified from ref. 23)



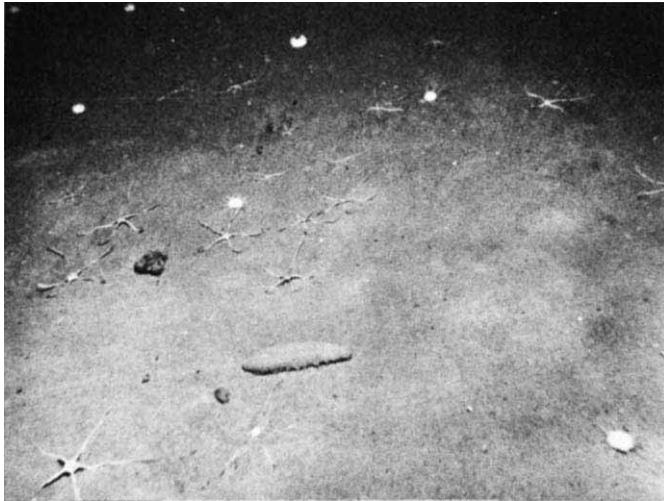


Fig. 4 About 3 m² of the sea floor at 1,800m depth SE of Massachusetts, USA. Alvin dive 685, 1976. Many brittle stars, *Ophiomusium lymani* (some in motion), white sea urchins, *Echinus affinis*, and one holothurian, Synallactidae. (Courtesy of Linda Morse, WHOI.)

inferior role. Brittle stars are very abundant in the more productive environments closer to the shore and are followed by cumaceans (another group of peracarid crustaceans). Bivalves dominate in the deep oligotrophic environment under the Sargasso Sea, followed by isopods. In Arctic Ocean sled samples from bathyal depths, sponges and polychaetes constitute nearly 25% each and crustaceans and ascidians 15% each¹⁸.

An evaluation of the composition of the hadal trench fauna cannot yet be based on quantitative data. Instead, Table 1 shows the average relative abundance of individuals collected in sled trawls in the four most thoroughly investigated trenches (which have an almost equal relative abundance). In these, and in all other trenches explored so far, the dominance of holothurians is as high as that of the polychaetes bathyally and in core/grab samples. Holothurians are followed in abundance by bivalves, polychaetes, and isopods. The preponderance of bivalves here and at 4,700 m in the north-western Atlantic (Table 1) may be due to winnowing during sampling at such great depths, resulting in loss of some of the lighter elements of the samples.

Two groups of protozoans of macrofaunal size, both agglutinating rhizopodes, are remarkably abundant, at least in some areas. The xenophyophores which are usually several centimetres in diameter may have pseudopodia extending 10–15 cm (ref. 20) and thereby be of great ecological importance. The other group is a hitherto overlooked, new superfamily, the Komokiacea (Fig. 2). These have a world-wide distribution, with a high relative abundance in hadal trenches and in the abyssal oligotrophic North Pacific. In the latter area their volume exceeds that of the metazoans and equals that of other foraminiferans²¹.

Extensive quantitative investigations in the Pacific²² show that polychaetes also constitute the dominant group at abyssal depths in biomass (up to 50%), although with the same decrease in importance in oligotrophic areas as was demonstrated when considering numerical abundance. Furthermore the biomass of crustaceans is relatively high in the oligotrophic areas (up to 25%) but almost obsolete in the productive areas. Similarly, echinoderms vary greatly in relative abundance and are nearly as dominant in weight in the south-western Pacific as the polychaetes. In the Arctic Ocean¹⁸, bivalves dominate in biomass (58%), and sponges (16%) are also more important than polychaetes.

The data of quantitative investigations demonstrate a pronounced decrease in total biomass with increasing depth

and distance from land, reflecting the reduced food input (Fig. 3.) Thus, 50 miles off California at about 200 m depth the total biomass is about 40 g m⁻², 50 miles further out, at 4,400 m, the biomass drops to 3.5 g m⁻²; a further decrease to 0.2–0.5 g m⁻² occurs 600 miles from the coast at about the same depth²³.

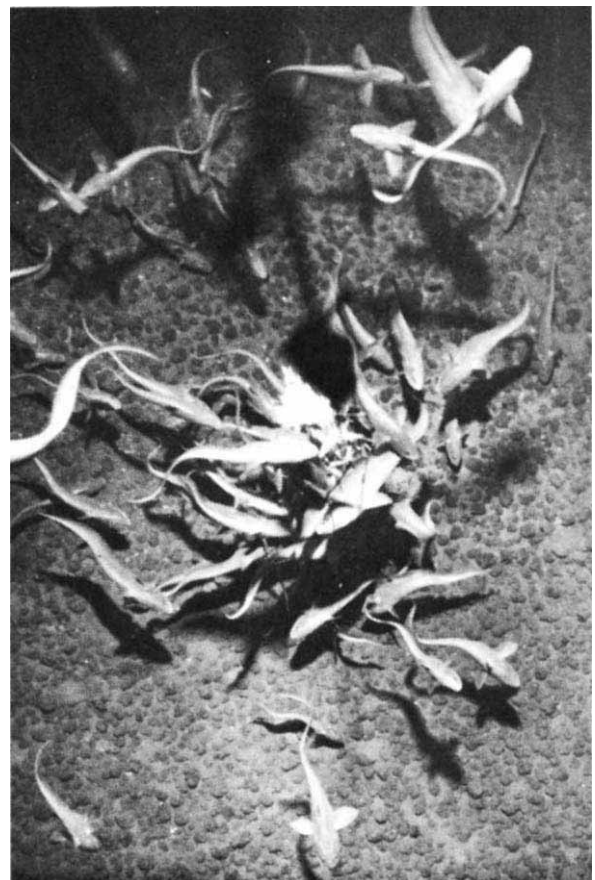
In samples consisting of many small, equal-sized animals, numerical density is a better measure for biological biomass than the ordinary wet weight measurements which include inorganic shells of bivalves and skeletons of echinoderms and crustaceans. Comparable numerical densities are available from the north-western Atlantic and show the same marked decline, from 21,000 individuals per m² at the edge of the continental shelf and between 8,600 and 1,700 on the slope to only 30–150 individuals at 5,000 m under the Sargasso Sea and 200–700 on the slope of the oceanic Bermuda island group¹.

Exceptions to the general rule of macrofaunal decrease with depth and distance from land are the Arctic and Antarctic Oceans and deeper water off upwelling areas with a particularly high production in the euphotic zone. A relatively rich fauna also exists in most trenches close to continents and larger islands, probably because the trenches act as traps for organic material. Biomass as high as 40 g m⁻² has been reported from the Aleutian Trench²².

The megafauna: Photographic documentation, combined with trawling or collecting from submersibles, are recently introduced methods which supplement conventional trawling, thus allowing a more detailed and reliable evaluation of the composition and distribution of the epibenthic megafauna.

Series of bottom photographs from bathyal depths in the north-western Atlantic off the Carolinas²⁴ show dominance

Fig. 5 Grenadiers (macrourids) and possibly brotulids attacking bait on an extremely oligotrophic bottom, covered with manganese nodules, at a depth of 5,900 m in the central North Pacific. The width of the picture equals about 2 m. (Courtesy of John D. Isaacs, SIO.)



of the quill worm *Hyalinoecia* (up to 15 individuals per m² at 450 m) and the brittle star *Ophiomusium* (2–6 per m² at 1,700–2,200 m). Although nearly half of the identifiable species are echinoderms, only one other brittle star is abundant. Both brittle stars have a wide vertical distribution, while most other megafaunal species show the same narrow zonation as macrofauna species⁴.

Ophiomusium and *Hyalinoecia* also dominate at similar depths off New England²⁵ (Fig. 4). Except for a sea urchin, the other 28 identifiable species (17 of which are fish) are rare. At about 2,000 m the megafaunal numerical density is three orders of magnitude less than the macrofaunal density in anchor dredge samples.

Calculation of the biomass of the megafauna components in trawl samples from the same area²⁶ showed that macrofaunal and megafaunal biomass are, however, of the same order of magnitude. The most abundant taxa are fish and echinoderms. Generally fish dominate in number and particularly in weight at all depths (140–1,900 m), and their size increases with increasing depth.

In recent years an extraordinary abundance of large, mobile fish and other vigorous scavengers has unexpectedly been demonstrated in the deep sea by means of baited ballast and a camera suspended above the bait. Every 10 min a still photograph or motion pictures are taken, and after 24–48 hours the camera is automatically released and returned (using floats) to the surface.

This free-vehicle research has been conducted at almost all depths in all main oceans^{8,27}. At bathyal depths the usual sequence of events is first the appearance of large shrimps, brittle stars, and amphipods, followed about 30 min later by a variety of fish which often suddenly scatter because a very large sleeper shark (*Somniosus*) moves in. Crabs, sea urchins, snails, and so on are responsible for the final cleaning.

At abyssal depths fish and amphipods are the overall dominating scavengers. Large grenadier fish (*Coryphaenoides* and other macrourids) are very active (Fig. 5), while eelpout-like fish (Lycodidae) and brotulids remain almost motionless for a very long time or at most nibble gently at the bait. Lysianassid amphipods up to 28 cm long (twice as long as any amphipod previously known) were photographed at 5,300 m in the North Pacific²⁸ and at 7,000 m in the Peru–Chile Trench (Fig. 6). At 9,600 m in

Fig. 6 Multitude of amphipods and several shrimp-like mysids gather around the bait at 7,000 m depth in the Peru–Chile Trench, stripping the bait in a few hours. (Courtesy of John D. Isaacs, SIO.)

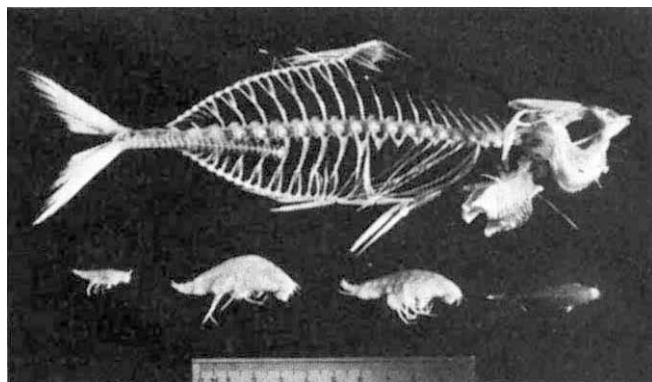


Fig. 7 A thoroughly skeletonised bait fish and four of thousands of large, ravenous amphipods from traps set at 9,600 m depth in the Philippine Trench.

the Philippine Trench thousands of amphipods were attracted to fish bait, and subsequently more than 10 litres were caught in baited traps (my experience in 1975 on board R/V Washington of Scripps Institution of Oceanography). All belonged to one species of Lysianassidae (Fig. 7), of which only four specimens had previously been collected in the trench²⁹.

It is remarkable that scavenging fish and amphipods are particularly abundant in the central North Pacific Gyre, where the benthic standing crop is extremely low^{11,22,23}. The ubiquitous scavengers must therefore rely on large 'food parcels' descending from above, for example, dead bodies of fish and marine mammals and food fragments from surface-feeding fish. Once on the bottom, the food is quickly located and consumed by opportunistic scavenging fish and amphipods.

Amphipods caught in baited traps in the North Pacific³⁰ and in the Philippine Trench were all without broodpouch plates or eggs in spite of their large size. This fact suggests that onset of reproductive maturity may be conditioned by the intake of a large meal³⁰ and may be postponed until one such is available.

Nothing is known about the density of scavengers or the distance from which they are attracted to food. Obviously, in order to sustain this roaming fauna, larger amounts of 'food parcels' must reach the bottom than previously anticipated. In view of the abundance of scavengers, the faeces they disperse over the sea floor is probably an important source of energy for deposit feeders⁸, thus contributing to the maintenance of life in the deep sea.

¹ Sanders, H. L., Hessler, R. R. & Hampson, G. R. *Deep-Sea Res.* 12, 845–867 (1965).

² Hessler, R. R. & Sanders, H. L. *Deep-Sea Res.* 14, 65–78 (1967).

³ Wolff, T. *Galathea Rep.* 5, 129–162 (1961).

⁴ Sanders, H. L. & Hessler, R. R. *Science* 163, 1419–1424 (1969).

⁵ Coull, B. C. *Mar. Biol.* 14, 48–51 (1972).

⁶ Sanders, H. L. *Ant. Nat.* 102, 243–282 (1968).

⁷ Sanders, H. L. *Brookhaven Symp. Biol.* No. 22, 71–81 (1969).

⁸ Dayton, P. K. & Hessler, R. R. *Deep-Sea Res.* 19, 199–208 (1972).

⁹ Grassle, J. F. & Sanders, H. L. *Deep-Sea Res.* 20, 643–659 (1973).

¹⁰ Turekian, K. K. et al. *Proc. natn. Acad. Sci. U.S.A.*, 72, 2829–2832 (1975).

¹¹ Hessler, R. R. & Jumars, P. A. *Deep-Sea Res.* 21, 185–209 (1974).

¹² Jumars, P. A. *Mar. Biol.* 30, 253–266 (1975).

¹³ Jumars, P. A. *J. mar. Res.* 34, 217–246 (1976).

¹⁴ Thiel, H. *Int. Revue Ges. Hydrobiol.* 60, 575–606 (1975).

¹⁵ Saidova, K. M. *Trudy Akad. Nauk S.S.S.R.* 86, 144–176 (1970).

¹⁶ Tietjen, J. H. *Deep-Sea Res.* 18, 941–957 (1971).

¹⁷ Coull, B. C. et al. *Mar. Biol.* 39, 233–240 (1977).

¹⁸ Paul, A. Z. & Menzies, R. J. *Mar. Biol.* 27, 251–262 (1974).

¹⁹ Wolff, T. *Deep-Sea Res.* 17, 983–1003 (1970).

²⁰ Lemche, H., Hansen, B., Madsen, F. J., Tendal, O. S. & Wolff, T. *Vidensk. Meddr dansk naturh. Foren.* 139, 263–336 (1976).

²¹ Tendal, O. S. & Hessler, R. R. *Galathea Rep.* 14 (in the press).

²² Filatova, Z. A. in: *The Pacific Ocean VII: Biology of the Pacific Ocean, Part 1: The Deep-sea Bottom Fauna*, 202–216 (1969) (Transl. U.S. Naval Oceanogr. Office 1970).

²³ Zenkevich, L. A., Filatova, Z. A., Belyaev, G. M., Lukyanova, T. S. & Suetova, I. A. *Byull. mosk. Obshch. Ispyt. Priro.* 76, 27–34 (1971).

²⁴ Rowe, G. T. & Menzies, R. J. *Deep-Sea Res.* 16, 531–537 (1969).

²⁵ Grassle, J. F., Sanders, H. L., Hessler, R. R., Rowe, G. T., & McLellan, T. *Deep-Sea Res.* 22, 457–481 (1975).

²⁶ Haedrich, R. L., Rowe, G. T. & Polloni, P. T. *J. mar. Res.* 33, 191–212 (1975).

²⁷ Isaacs, J. D. & Schwartzlose, R. A. *Scient. Am.* 233, 85–91 (1975).

²⁸ Hessler, R. R. & Mills, E. L. *Science* 175, 636–637 (1972).

²⁹ Dahl, E. *Galathea Rep.* 1, 211–241 (1959).

³⁰ Schulenberg, E. & Hessler, R. R. *Mar. Biol.* 28, 185–187 (1974).

Drugs from the sea—fact or fantasy?

P. T. Grant & A. M. Mackie

NERC, Institute of Marine Biochemistry, Aberdeen, UK

Drugs from the sea are as much a potential marine resource as cultivated fish, and mineral deposits. The study of the chemical structure and properties of unusual metabolic products of marine life is a subject where marine ecology and the experimental sciences of chemistry, biochemistry, pharmacology and medicine share a common and complementary interest. The development of ad hoc collaboration between specialists has advanced basic knowledge and resulted in a significant feedback to marine biology and ecology as well as in the development of some useful drugs.

OUR present armamentarium of drugs used to treat man and his domestic animals is mainly comprised of either natural products of unusual chemical structure and function derived from terrestrial microorganisms, plants and animals, or synthetic analogues of these products. Evolutionary adaptation has presumably resulted in metabolic diversity and cellular specialisation for their biosynthesis. In some cases it has been shown that these bioactive products of metabolism represent one mechanism of species survival in the highly competitive situations that can arise in natural ecosystems¹.

The clinical success of sulphonamides (Prontosil) and penicillin in the period 1938–45, led to a general acceptance of the idea that many human diseases would be amenable to treatment with selectively toxic chemicals. This led to the subsequent and successful search for new drugs. There is now some concern about the diminishing returns resulting from a continued emphasis on the screening of terrestrial life forms and a developing interest in the certainty that substances of therapeutic importance are to be found in the vast and diverse range of marine life—an estimated 500,000 species in 30 phyla. This certainty arises from the numerous reports of the pharmacological and toxicological effects elicited by extracts of marine organisms and from our knowledge of substances of ecological importance, such as the potent venoms and poisons concerned with offence and defence reactions. There is also a very extensive folklore and old records of a more or less anecdotal nature that are relevant to both human medicine and marine ecology^{2–4}. In this article, actual and potential therapeutic agents have been broadly classified as either 'antibiotics' that may be used to selectively control and destroy pathogens and parasites that cause disease, or systemic drugs that can modify or control body metabolism and function.

Antibiotics

Marine bacteria, fungi, phytoplanktonic and higher algae, as well as some invertebrates have been shown to contain antibacterial, antifungal and antiviral activities. The extensive literature on this subject has been reviewed^{2,3,5}. Some seasonal variation has been recorded. This would imply that the physical and chemical conditions of the microenvironment are critical for the expression of antibiotic activity by these organisms—a conclusion consistent with our extensive knowledge of the culture conditions necessary for the production of classical antibiotics such as penicillin.

In a study of the antibiotic activity of microorganisms isolated from various depths of the oceans (0–3,500 m), Krasil'nikova⁶ found that 124 of the 326 cultures collected were active against the test organisms *Staphylococcus aureus*, *Escherichia coli*, *Myobacterium luteum* and *Saccharomyces cerevisiae*. One actinomycete was also collected with a broad antibacterial spectrum. It is therefore not surprising that a variety of antibiotic activities have been observed in samples of filtered seawater³. At the same time relatively few studies have been taken to the point of isolation of the active principles involved. It is difficult to evaluate the relevance of most of these studies because the ratio of the minimum curative dose of the pure substance to the maximum tolerated dose for the infected animal is the first criterion used to determine the therapeutic potential of a substance with antibiotic activity. Moreover, in some cases, the active principle has turned out to be mainly of ecological significance, such as acrylic acid or brominated phenolic derivatives. One classical example of the potential of marine organisms has its origins in the observation that the culture fluid of a fungus, isolated from a sewage outfall into the sea, was toxic to a wide range of bacteria with an antibacterial spectrum similar to that of penicillin. Subsequently, Abraham, Newton and Florey showed that the fungus produced a number of antibiotics, notably a molecule containing a fused β -lactam-dehydrothiazine ring system (termed Cephalosporin C), as well as one with a fused β -lactam-thiazolidine ring system characteristic of the penicillin molecule⁷. Cephalosporin C and chemically-modified derivatives were toxic to a wide range of pathogenic bacteria, resistant to inactivation by the enzyme penicillin- β -lactamase and of low toxicity to man. Resistance to the inactivating enzyme was a very desirable property because a feature of therapy with penicillin was the development of strains of bacteria resistant to penicillin through selection of strains that produced the enzyme.

Other defined chemicals of marine origin with selectively toxic properties have been used in clinical medicine. Thelapin (a derivative of coumaran-2-spiro-cyclohexane, containing three bromine atoms per molecule) obtained from extracts of the polychaete, *Thelaps setosus*, is an antifungal agent with a potency comparable to griseofulvin (a derivative of a coumaran-2-spiro-cyclohexane containing one chlorine atom per molecule), to which it bears a strong structural resemblance⁸. Kainic acid (2-carboxy,3-carboxymethyl,4-isopropenylpyrrolidine) is an efficient anthelmintic used in Japan to eliminate intestinal worms. Kainic acid is the active principle of the red algae, *Digenia simplex*⁹. Dried powders or syrups prepared from extracts of *Digenia* and other red algae have long been used as anthelmintics, with low toxicity to man, in Asia and in some countries bordering the Mediterranean³. The chemical nature and properties of antibiotics produced by marine organisms may well be the most rewarding area for definitive study in terms of human medicine. A programme of culture, screening, the selection of high-yielding strains and other development would require a considerable and broad-based effort, even if carried out in association with pharmaceutical companies who have experience with similar work involving terrestrial organisms. Apart from the speculative nature of such a venture, a purely mercenary aspect could limit support—as soon as the chemical structure of a drug is known an attempt can be made to make it available by

total chemical synthesis. While this can happen in some cases, most of our antibiotics such as the penicillins, streptomycin and tetracyclines are extremely complex structures and continue to be produced by the culture of high-yielding strains of microorganisms on a commercial scale.

Systemic drugs

There are about 2,000 venomous or poisonous species widely distributed among the marine fauna, with some concentration in tropical seas. The property of venom or poison biosynthesis is one mechanism that ensures species survival in densely populated ecosystems and represents a very successful form of chemical evolution and metabolic diversity from the common pattern. The general biology, chemistry and possible treatment of human illness caused by venom and by eating poison-laden fish, has been comprehensively reviewed by Halstead¹⁰. One of the most potent poisons known is tetrodotoxin (an amino perhydroquinazoline), a constituent of the liver and gonads of certain puffer fish (Tetraodon group)¹¹, porcupine fish (Diodontidae), ocean sunfish (Molidae), newts (Taricha) and some South American frogs. The flesh of the puffer fish is regarded as a great delicacy by Japanese gourmets, and a chef's reputation depends entirely on his ability to remove every trace of the toxic organs from the fish, and on the number of his guests who remain alive after eating his food. The dramatic and toxic effects of sampling sunfish roe is recorded by Captain Cook in his diary for 1774 (ref. 31). A lethal dose to man is about 10 µg per kg body weight and tetrodotoxin is about 160,000 times more potent in blocking axonal conduction than cocaine¹². It is extensively used as an inhibitor in neurophysiological research since it has a unique action on the gating of sodium channels or holes in excitable membranes. At low dosage levels, tetrodotoxin has been used clinically as a muscle relaxant and a pain killer in neurogenic leprosy and in terminal cancer⁵. It has also potential as a local anaesthetic¹³. The conclusion that tetrodotoxin is an endogenous fish poison¹¹ requires direct proof. Saxitoxin (a trialkyl tetrahydropurine derivative), is produced by strains of dinoflagellates of the genus *Gonyaulax* and has very similar but not identical pharmacological properties¹⁴. A massive overgrowth of these dinoflagellates and other phytoplankton can occur at certain times of the year in conditions that are not fully understood. This overgrowth develops to the point where the sea becomes discoloured and is termed a red-tide phenomenon. The bloom of phytoplankton can result in mass mortalities of fish due to toxins or to oxygen deprivation in the area of the bloom—but not of shellfish that consume the organisms and accumulate the toxin. This food-chain mechanism is presumably responsible for the periodic outbreaks of mussel poisoning or paralytic shellfish poisoning of man that have been recorded in recent years and in historical records¹⁴. It remains to be established whether saxitoxin is produced together with other toxins by the genus *Gymnodinium* which is also associated with red-tide phenomena. Apart from the venoms and poisons there are many reports that extracts of marine organisms have pharmacological effects on mammalian and other tissue preparations which can be described in such terms as anti-tumour, hellucogenic, neuroactive, cardioactive and cardiovascular^{2,3,5}. In some cases, the active principle has been identified. Interest in the ability of several genera of octopus to paralyse their prey by means of secretions from the posterior salivary glands led to the isolation of the peptide with the linear sequence, Pyroglu-Pro-Ser-Lys-Asp-Ala-Phe-Ile-Gly-Leu-Met-NH₂. This peptide from the salivary gland of *Eledone moschata* was termed eledoisin¹⁵, and proved to be a highly potent hypotensive agent in dogs—at least 50 times as active as histamine, acetylcholine and bradykinin on an equivalent

basis. Eledoisin and the related hexapeptide, Lys-Phe-Ile-Gly-Leu-Met-NH₂ (ref. 16), also stimulate extravascular smooth muscle and effect an atropine-resistant increase in salivary secretions. A cardioactive chemical, possibly an aromatic amine³² has been isolated in crystalline form from the branchial heart tissue of the primitive hagfish, *Eptatretus stoutii*¹⁷. The amine is a potent stimulator and pacemaker of the mammalian heart. This observation may be related to the fact that the heart of the hagfish consists of four pulsatile organs that operate independently—the systemic or branchial organ being aneural.

The possibility that some of the unusual metabolic products of marine organisms would have a selectively toxic action on tumour tissues in mammals remains an interesting possibility. Several have been identified, such as mercenene, a glycopeptide present in the common clam, *Mercenaria mercenaria*, during the summer months which is apparently non-toxic to mice at therapeutic levels¹⁸, and ill-defined substances present in the tentacles of two topical polychaetes¹⁹. The latter study would partly sustain the claims of folklore that cooked tentacles of *Lanice conchilega* had beneficial effects on certain forms of human cancer. One potent clinical anti-tumour agent has its origins in the discovery by Bergmann in 1968 that the West Indian sponge, *Tethya crypta*, contained large amounts of unusual arabinosyl-nucleosides. This led to the chemical synthesis and testing of a variety of nucleosides containing the unusual sugar moiety²⁰. 1-β-D-arabinofuranosyl cytosine (Ara-C) is a potent inhibitor of mammalian cell growth and is used extensively in chemotherapy in cancers such as leukaemias^{21,22}. Ara-C is known to be slowly phosphorylated to the 5'-triphosphate and has been shown to be incorporated into RNA and DNA and to inhibit DNA synthesis. Recently, a liponucleotide analogue, arabinofuranosylcytosine-5-diphosphate-1,2-dipalmitin has been synthesised and shown to be a potent inhibitor of leukemic cells (strain L 5178Y) in mice²³. It is claimed that cellular metabolism of the liponucleotide would generate Ara-CMP and hence the 5'-triphosphate at a faster rate than that resulting from direct phosphorylation of Ara-C by tissue kinases.

Attention has so far been concentrated on the nature and possible functions of the products of metabolic diversity and cellular specialisation in poikilothermic (cold blooded) marine organisms. Many patterns of metabolism are, however, common to all forms of life and certain hormones and other agents of metabolic control have a potential role in certain human disease states. In most cases these substances have been obtained from lower vertebrates sharing the same ultimate primitive ancestors as the mammals. One example was the extensive use of fish insulin, isolated mainly from the discrete islets of Langerhans of the tuna fish, to treat diabetics in Japan in the period 1939–45 (ref. 24). Other fish insulins, such as that of the cod *Gadus callarias*, have essentially the same specific hormonal activity in mammals as that of the homologous insulin. About 40% of the amino acid residues in cod insulin differ from those of bovine or porcine insulin but the invariant residues and features necessary for hormonal activity are conserved²⁵. In contrast, it was shown by Copp and others that the peptide hormone calcitonin, isolated from the ultimobranchial glands of the salmon (*Oncorhynchus*), was about 100–200-fold more active in man than the corresponding porcine hormone. This dramatic difference in activity has now been attributed to the greater resistance of the salmon hormone to degradation in the circulation of the mammal²⁶. A rather surprising development in recent years has been the discovery of relatively large amounts of prostaglandins in strains of the West Indian gorgonian, *Plexaura homomalla*²⁷, and in other marine organisms²⁸. The prostaglandins and thromboxanes constitute a family of related compounds that can be regarded as derivatives of the twenty-carbon atom monocyclic fatty acid, prostanoic acid.

These compounds were originally identified in mammalian tissue by Bergström in 1967 and have subsequently been implicated as the regulators of a very wide range of critical life processes in mammals, and in some disease states. Only trace amounts occur in the blood and tissues, both the prostaglandins and the thromboxanes being formed by controlled biosynthesis from essential fatty acids through the intermediate prostaglandin endoperoxide. For this reason the gorgonians may well prove to be a useful starting material for chemical modification of prostaglandins to the mammalian types. This route may be a useful alternative to the successful but more complex process of total synthesis. An interesting feedback to marine biology is some evidence that the induction of prostaglandin synthesis (by hydrogen peroxide) is responsible for synchronous spawning in both male and female molluscs²⁹. It is thus tempting to speculate that the chemical signal effecting epidemic or synchronous spawning in the natural environment may be the concentration of oxidising free radicals generated by the light-induced photolysis of water.

Postscript

The only integrated proposals for the definitive study of a selection from the vast number of biologically active substances of marine origin have been oriented towards human medicine. A report to the Congress of the United States³⁰, contained a recommendation that a National Institute of Marine Medicine and Pharmacology should be established

as one of the National Institutes of Health. This recommendation has not so far been implemented, but a pharmaceutical company has recently opened a new laboratory in New South Wales, Australia—the Roche Research Institute of Marine Pharmacology.

- ¹ Whittaker, R. H. & Feeny, P. P. *Science* **171**, 757 (1971).
- ² der Marderosian, A. J. *Pharm. Sci.* **58**, 1 (1969).
- ³ Baslow, M. H. *Marine Pharmacology* (William and Wilkins, Baltimore, 1969).
- ⁴ Russell, F. E. *Adv. mar. Biol.* **3**, 255 (1965).
- ⁵ Ruggieri, G. D. *Science* **194**, 491 (1976).
- ⁶ Krasil'nikova, E. N. *Microbiology* **30**, 545 (1962).
- ⁷ Abraham, E. P. & Newton, G. G. F. *Adv. Chemother.* **2**, 23 (1965).
- ⁸ Higa, T. & Scheuer, P. J. *Tetrahedron* **31**, 2379 (1975).
- ⁹ Ueno, Y. *et al. Proc. Japan Acad.* **33**, 53 (1957).
- ¹⁰ Halstead, B. W. *Poisonous and Venomous Animals of the World* 1-3 (U.S. Government Printing Office, Washington, DC, 1965).
- ¹¹ Larson, E. *et al. in Bioactive Compounds from the Sea* (eds Humm, H. J. & Lane, C. E.) 139-141 (Marcel Dekker, New York, 1974).
- ¹² Kao, C. Y. *in Bioactive Compounds from the Sea* (eds Humm, H. J. & Lane, C. E.) 115-121 (Marcel Dekker, New York, 1974).
- ¹³ Ogura, Y. & Mori, Y. *Eur. J. Pharmac.* **3**, 58 (1968).
- ¹⁴ Alam, M. *Tex. Rep. biol. Med.* **33**, 183 (1975).
- ¹⁵ Anastasi, A. & Erspamer, V. *Archs Biochem. Biophys.* **101**, 56 (1963).
- ¹⁶ Bergmann, J. *et al. Experientia* **30**, 1315 (1974).
- ¹⁷ Jensen, D. *Comp. Biochem. Physiol.* **10**, 129 (1963).
- ¹⁸ Li, C. P., Goldin, A. & Hartwell, J. L. *Cancer Chemother. Rep.* **2**, 4, 97 (1974).
- ¹⁹ Tabrah, F. L., Kashiwagi, M. & Norton, T. R. *Science* **170**, 181 (1970).
- ²⁰ Evans, J. S. *et al. Proc. Soc. exp. Biol. Med.* **106**, 350 (1961).
- ²¹ Loo, T. L. *et al. Ann. N. Y. Acad. Sci.* **255**, 252 (1975).
- ²² Kramer, W. D. *Ann. Intern. Med.* **82**, 684 (1975).
- ²³ Rietz, C. R. H. *et al. Science* **196**, 303 (1977).
- ²⁴ Nagasawa, K. *J. Ass. off. agric. Chem.* **51**, 326 (1968).
- ²⁵ Reid, K. B. M., Grant, P. T. & Youngson, A. *Biochem. J.* **110**, 289 (1968).
- ²⁶ Queener, S. F. & Bell, N. H. *Metabolism* **24**, 555 (1975).
- ²⁷ Weinheimer, A. J. *in Bioactive Compounds from the Sea* (eds Humm, H. J. & Lane, C. E.) 231-238 (Marcel Dekker, New York, 1974).
- ²⁸ Nomura, T. & Ogata, H. *Biochim. biophys. Acta.* **431**, 127 (1976).
- ²⁹ Morse, E. *et al. Science* **196**, 298 (1977).
- ³⁰ *Our Nation and the Sea, Report of the Commission on Marine Science, Engineering and Resources*: 119-121, (U.S. Government Office, Washington, DC, 1969).
- ³¹ Beaglehole, J. C. *The Journals of Captain James Cook*, **2**, 534 (Cambridge University Press, Cambridge, 1961).
- ³² Pollack, G. H. *Science*, **196**, 731 (1977).

Bioluminescence of marine organisms

Peter J. Herring

Institute of Oceanographic Sciences, Wormley, Godalming, Surrey, UK

Luminescence is a feature common to many marine organisms. Certain squid and fish utilise luminous bacteria as symbiotic sources of light, and the symbionts have physiological features which are of potential advantage in the association. The luciferins of some marine fishes, squids, crustaceans and coelenterates are of very similar chemical structure, though differing markedly from the luciferins of non-marine forms. Luminescence in many marine animals is probably used for ventral camouflage purposes.

THOSE whose lives have been closely associated with the sea and its resources have long been aware of both its general luminescence and that of many of its inhabitants. The first written record concerning luminescence of the sea has been attributed to Anaximenes (about 500 BC) and accounts of luminous marine animals have been available at least since the younger Pliny described the luminescence of *Pholas* and medusae¹. Bioluminescence has always been a source of wonder, often of awe, accentuated in earlier years by the conceptual difficulties of separating the phenomenon of light from that of heat. Indeed there is still occasional (unfounded) alarm that luminescent 'sparks' may trigger off explosions in oil tankers. By the mid-nineteenth century a variety of deep water marine animals had been taken which bore structures variously described as pores, slime glands or accessory eyes. It only became apparent that these

organs were for the production of light, rather than its perception, when the naturalists aboard HMS Challenger observed the emission of light from the ventral photophores of euphausiids and hatchetfishes². Only in this last century has the nature and occurrence of bioluminescence begun to be elucidated. Taxonomists have even occasionally enshrined the real (or sometimes imagined!) luminous abilities of many marine organisms in their names. Examples include *Photobacterium phosphoreum* (bacterium) *Pyrocystis* (dinoflagellate) *Pelagia noctiluca* (medusa) *Euphausia*, *Pyrocypis*, *Metridia lucens* (crustaceans) *Watasenia scintillans*, *Symplectoteuthis luminosa* (squids) *Pyrosoma* (tunicate) *Lampanyctus*, *Photoblepharon* and *Benthosema* (fishes). Even in common parlance lanternfishes and fireworms provide marine counterparts to lightning bugs and fireflies.

This review deals briefly with three aspects of marine luminescence, namely the symbiotic luminous bacteria, the comparative chemistry of marine systems, and the use of luminescence as a camouflage.

Symbiotic luminous bacteria

All luminous bacteria require marine or other saline conditions, and they are present at all depths in the ocean, albeit in relatively small numbers. They can be grown in mass cultures in the laboratory and consequently more is known about the chemistry and molecular control of their luminescent system than that of almost any other organism^{3,4}. The luminescence of bacteria is normally continuous, with maximum intensity corresponding to a similar

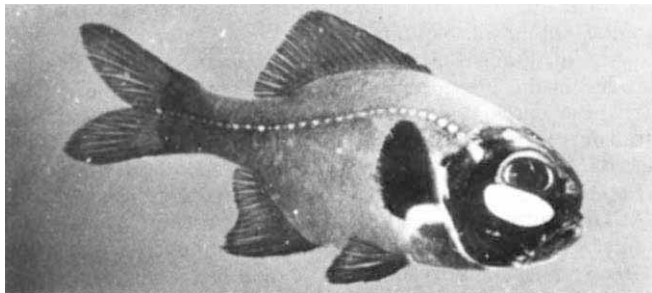


Fig. 1 *Photoblepharon palpebratus*—the large light organ below the eye contains luminous bacteria.

maximum in luciferase synthesis. The initiation of luciferase synthesis is dependent on the accumulation of an auto-inducer in the culture medium; very dilute cultures do not luminesce¹. Some cephalopods and many fishes have taken advantage of the availability of luminous marine bacteria and developed organs in which a population of these microorganisms is maintained, thereby providing the host with a symbiotic source of light. Two genera of luminous bacteria, *Photobacterium* and *Beneckeia*, are currently recognised⁵. It seems that *Beneckeia* includes only free-living strains, whereas all symbiotic strains should be included in *Photobacterium*⁵. It has recently been shown that the synthesis of the luminous system of some species of *Photobacterium* is enhanced at growth-limiting low oxygen concentrations⁶. This has obvious ecological significance in the symbiont's environment, as do other features of its physiology⁷, and is not the case in *Beneckeia*⁶.

Some 60 yr ago enthusiasm for symbiotic bacteria as the source of animal light reached such a pitch that some investigators held it likely that all luminous organs contained symbionts. Symbiotic bacteria are now recognised as responsible for the luminescence of certain loliginid and sepiolid squids, and a considerable number of fishes. Most coastal luminous fishes utilise bacteria (Leiognathidae, Eretmophoridae, Acropomatidae, Steindachneriidae, Trachichthyidae, Monocentridae, Anomalopidae and some Apogonidae⁸). Bacterial systems are, however, also present in oceanic fishes at all depths, namely in the benthopelagic Macrouridae⁹, the mesopelagic Opisthoproctidae^{10,11} and the bathypelagic anglerfishes^{12,13}. In most fishes the luminous bacteria are associated with one or two gut diverticula. In the Anomalopidae and Monocentridae, however, each fish has a pair of bacterial organs, just beneath the eyes in the former (Fig. 1) and on the lower jaw in the latter. The other exceptions are the anglerfishes, in which bacteria are maintained in the esca, a luminous bulb placed at the tip of a modified dorsal fin ray. In the anglerfishes *Ceratias* and *Cryptopsaras* two or three more posterior fin rays are also modified into bacterial bulbs (the caruncles)¹³. All bacterial organs open to the exterior, either directly through a pore or through the gut lumen (Table 1). Although luminous bacteria from gut-associated organs are generally easy to culture on artificial media, self-luminous cultures have not yet been routinely obtained from either the Anomalopidae or the anglerfishes. Starvation of *Anomalops* leads to a gradual dimming and final extinction of its luminescence, correlated with a great reduction in the number of bacteria in the light organs¹⁴.

Animals utilising bacteria have particular problems with which to contend. The bacteria must be restricted to the particular organ, and not allowed to spread indiscriminately (luminous sand-hoppers suffering from a general uncontrolled infection of luminous bacteria are occasionally reported¹⁵). The culture must be maintained in good condition and without contamination by other bacteria. Some symbiotic strains produce antibiotic substances which may

inhibit contaminants¹⁶. Surplus or dead bacteria must also be removed. Most importantly the light organs of successive generations must be reinfected with the appropriate strain. While this presents little difficulty in gut-associated systems, the mechanisms used by other fishes, particularly the anglerfishes, are quite unknown. This may perhaps be a contributory reason for the very small number of light organs present in animals utilising bacteria. Bacteria luminesce continually, and some means of controlling the light emission from the organ is therefore necessary. This may be effected by obscuring it with chromatophores¹⁷, rotating the organ, occluding it with a shutter^{18,19} or by restricting its blood supply¹³.

In non-bacterial or intrinsic luminescent systems different organs tend to be used for different purposes. The very limited number of light organs present in animals using bacteria has resulted in a remarkable behavioural versatility in their use; the same organ may be used for prey illumination, escape from predators, and both schooling and sexual communication²⁰.

Comparative biochemistry of marine luminescence

The discovery in the seventeenth century of inorganic chemiluminescent systems, or 'phosphori', overlapped Boyle's classical investigations of the luminescence of rotting wood and meat, and for a long period subsequently all natural luminescence was generically ascribed to a 'phosphorus'. As late as the nineteenth century when Edward Forbes persuaded his friend George Wilson, the eminent chemist, to investigate the luminescence of the sea-pen *Pennatula*, Wilson concluded: "On the whole I believe it most probable that the animal secretes a spontaneously inflammable substance. It may be a compound of phosphorus, but it is not necessary to assume that it is"²¹. Organic chemiluminescent compounds were only discovered in the latter part of the nineteenth century, and were a necessary prelude to the understanding of the chemical nature of bioluminescence.

In 1887 Dubois demonstrated the luciferin-luciferase reaction on material from the mollusc *Pholas*²¹. It was subsequently found that this general reaction could be demonstrated in only a minority of species tested²². This led to the conclusion that a very large number of quite separate chemical systems had been evolved in different species. Even now the chemistry of very few luminous systems is understood in any detail^{23,24}. This to a considerable extent reflects the difficulties of obtaining adequate quantities of active material for investigation. For example, some 10,000 specimens of the medusa *Aequorea* were needed to produce 5 mg of the active protein²⁵, and 10,000 specimens of the squid *Watasenia* were also used in the initial isolation of its luciferin²⁶. Extracted chemical systems were, until very recently, apparently either of the luciferin-luciferase type or photoproteins. The latter are high molecular weight compounds which require no additional enzyme and emit light in the presence of oxygen and/or an

Table 1 Comparative features of bacterial and intrinsic light organs

Bacterial organs	Intrinsic organs
Restricted to a few squids and fishes	Widely distributed
Usually 1 or 2 organs (never > 4) per species	Often numerous (several thousand in some fishes)
Only one type in each species	May be several types in each species
Always open to the exterior	Often closed
Often associated with the gut	Not usually associated with the gut

appropriate trigger²⁴. In the case of aequorin, the most studied photoprotein²⁷, only calcium ions are necessary, and light emission will occur in the absence of oxygen. In other photoproteins, such as that from the polychaete *Chaetopterus*, molecular oxygen and other cofactors are necessary²⁸. The demonstration that either a luciferin-luciferase system or a calcium-triggered protein can be extracted from the sea pansy *Renilla*^{29,30}, depending on the methods used, and the fact that the 'chromophore' of aequorin³¹ is chemically very similar, if not identical, to *Renilla* luciferin³², has led to a reappraisal of the photoprotein concept³³. It now seems that photoproteins may be regarded as stable luciferin-luciferase complexes. If this complex remains stable when oxygenated then a 'pre-charged'³⁴ photoprotein such as aequorin is obtained³⁵. In general it is likely that the oxidation pathway proceeds through the formation of either a dioxetane^{31,23} or a linear peroxide³⁶.

The structure of the luciferin of the ostracod *Cypridina* was one of the first to be elucidated. It is an imidazopyrazine³⁷, and in recent years strikingly similar chemical structures have been found in a number of other marine luciferins. The luciferin of *Renilla* (and the chromophore of *Aequorea*) is a similar compound, as is that of the squid *Watasenia*³⁸, and the decapod shrimps *Oplophorus* and *Heterocarpus*^{38,39}. It is clear that these luciferins can be produced by the cyclisation of a simple tripeptide (composed of tryptophan, isoleucine and arginine in *Cypridina*), and a number of active model compounds have been synthesised using different combinations of amino acids⁴⁰.

A luciferin apparently identical to that of *Cypridina* has been found in a number of fishes including *Parapriacanthus*⁴¹, *Apogon*⁴², the midshipman *Porichthys*⁴³ and the lanternfish *Diaphus*⁴⁴. Ostracods were found in the gut of some *Parapriacanthus* specimens and as the fish stores its luciferin in gut diverticula it was suggested that the luciferin might be obtained from its crustacean food^{45,46}. Some support for this conjecture has come from recent work on *Porichthys notatus*. Californian populations of this fish are luminous, but Puget Sound populations are not—although they possess a full complement of photophores. They lack luciferin, but produce luciferase normally. When fed *Cypridina*, or injected with natural or synthetic *Cypridina* luciferin, they too luminesce normally⁴⁷⁻⁴⁹. This remarkable situation suggests the need for an exogenous source of luciferin for normal luminescence, or at least for its induction.

There is clearly a common link in the chemical nature of the luciferins of a number of taxonomically unrelated marine animals. Despite the fact that the luciferins of fireflies⁵⁰, the mollusc *Latia*⁵¹, the oligochaete *Diplocardia*⁵² and luminous bacteria have very different structures, it is nevertheless remarkable that luciferins with very similar structures have been obtained from four different marine phyla. It will be of great interest in the future to determine whether other marine species, many of whose systems have a superficial biochemistry rather different to that of *Cypridina* or *Renilla*, utilise luciferins of similar general structure.

The colour of the emitted light may be determined in a number of ways⁵³. Pioneer work on insects⁵⁴ and ostracods⁵⁵ showed that the colour of the light emitted during the oxidation of a luciferin is determined by the source of the luciferase. Changes in the molecular environment of the luciferin, whether brought about by attachment to the luciferase or by physical factors such as heat, cations or pH, can produce different emission wavelengths. Work with model luciferins of the *Renilla* type has shown that different degrees of ionisation, in turn partly dependent upon the particular tripeptide used, can give rise to very different emission maxima. It is therefore theoretically possible that different marine luciferins, derived from dif-

ferent tripeptides, could account for the observed variations in the colour of bioluminescence²³.

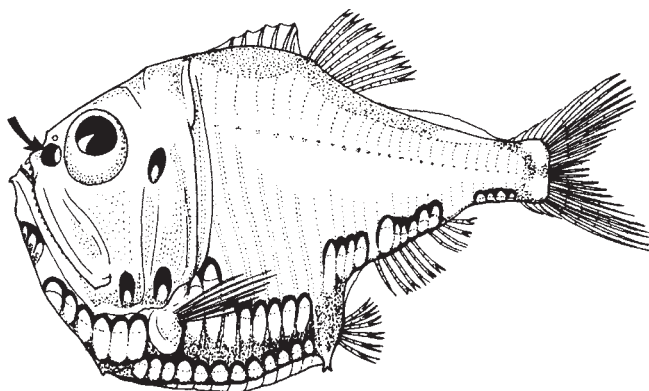
An intriguing feature of the luminous systems of some species is the presence of secondary emitters in the tissues. In these cases the energy of the luminescent reaction is transferred to a fluorescent material which then emits light at its own characteristic wavelength. Both *Renilla* and *Aequorea* contain a green fluorescent protein associated with the luminous tissue. *In vivo* light is green ($\lambda_{\max}$ 509 nm) but if the green fluorescent protein is removed a blue ($\lambda_{\max}$ 470–490 nm) *in vitro* emission is obtained. Addition of the green fluorescer to the *in vitro* system produces a green emission once more. Species lacking the green fluorescer (for example *Pelagia*, *Mnemiopsis*) emit only blue light both *in vivo* and *in vitro*⁵⁶⁻⁵⁸. Energy transfer has also been implicated in some fish⁴⁹ and squid³⁸ and may be responsible for bimodal emission spectra. It is tempting to consider that the red fluorescent proteins extracted from within the red-emitting photophores of the deep-sea fishes *Pachystomias Malacosteus* and *Aristostomias* may be involved in a similar process (P.J.H., unpublished).

Ventral camouflage

During the deep-sea expeditions of the late nineteenth and early twentieth centuries it became increasingly obvious that a great many deep sea animals bore light organs which were predominantly ventrally directed (Figs 2, 3 and 4). This is the case both in decapod crustaceans and in cephalopods, but is most prominently developed in fishes, both cartilaginous and bony. In 1916 Dahlgren, discussing cephalopods, noted "any animal situated below such a squid and looking upward at it from below would see a bluish light that would blend with the sunlight"⁵⁹. Rauther applied the same hypothesis to fish⁶⁰, but it was not until the idea of ventral camouflage was dealt with in detail by Fraser⁶¹ and Clarke⁶² that it achieved wide acceptance. In essence, animals living at depths still subject to background illumination from the surface are believed to illuminate their ventral surfaces so as to match the background, and eliminate the silhouette which would otherwise expose them to a predator from below. The use of this concept of ventral camouflage, or counter-illumination, was considered during World War 2 as a method of camouflaging planes during daylight sorties. Lights were to be fixed to their undersides to reduce or eliminate their silhouette, but the power requirement ruled out the idea at an early stage.

Recent work has been aimed at providing experimental support for this attractive hypothesis, but much of the

Fig. 2 The hatchetfish *Argyropspectus aculeatus*. The photophores are directed ventrally, with the exception of the preorbital organ (arrowed) which shines into the eye. This organ may act as a reference standard for comparison with ambient light.



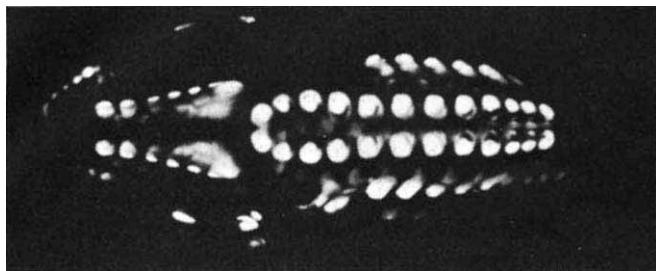
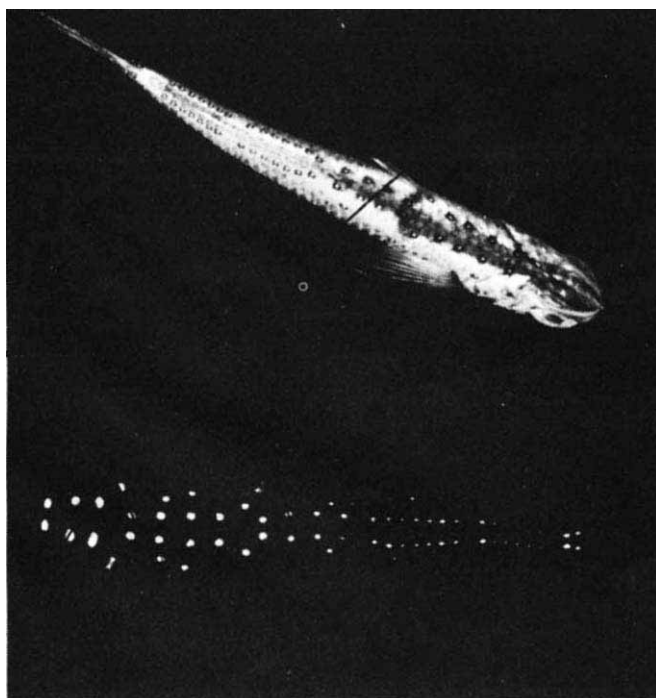


Fig. 3 Ventral view of a luminescing *Argyropelecus olfersi*.

evidence is still circumstantial. Counter-illumination provides a rationale for the otherwise anomalous positions of numerous photophores. Many otherwise transparent squids have photophores beneath the eyes and ink sac, their only opaque organs. Many crustaceans also have photophores beneath the eyes. In decapods such as *Oplophorus* the photophores on the pereopods will contribute to the ventral illumination of the body when these limbs are folded in a swimming position. Ventral camouflage provides an explanation for the extraordinary anatomy of the mesopelagic fish *Opisthoproctus* (Figs 5, 6). The light from its single bacterial rectal organ is spread over the whole flattened sole of the fish by an elaborate light guide system^{10,11}. Similar anatomical diffusion of the light over the ventral surface is found in many other fishes utilising bacteria. The restricted number of light organs in such animals requires that elaborate accessory optical structures be present if a general ventral illumination is to be achieved^{17,64}.

Detailed anatomical and experimental studies of hatchetfish photophores have shown that their reflecting systems are so arranged that the angular distribution of the emitted light almost exactly matches that of the ambient downwelling daylight (Fig. 7)⁶⁵⁻⁶⁷. This arrangement ensures that the ventral camouflage is equally effective when observed from almost any angle. Hatchetfishes have a filter pigment, with a transmission maximum at about 480 nm, in the aperture of their photophores⁶⁸. Similar filter pigments are

Fig. 4 A lanternfish (*Myctophum* sp.) showing the ventral array of photophores (a) and its appearance when luminescing (b).



found in many other fishes, squids and crustaceans (Fig. 8). Measurements of the emission spectra of the photophores of certain fish before and after the removal of the filter pigments, have shown that the effect of the filter is to modify the emission spectrum so that it finally closely approximates that of the down-welling daylight (Denton, E.J., and Herring, P.J., unpublished). Some animals, at least, are therefore able to match both the angular and spectral distribution of ambient daylight. Such a system will obviously be most effective if the fish remains in a constant light intensity. It has indeed been demonstrated that some animals, in the form of scattering layers, remain very close to particular isolumens⁶⁹. Changes in ambient intensity also affect the depth distribution of many species, not only by day but also by night⁷⁰. If, on the other hand, an animal can alter the intensity of its own bioluminescence it achieves much greater behavioural flexibility, as its movements in the water column can then become independent

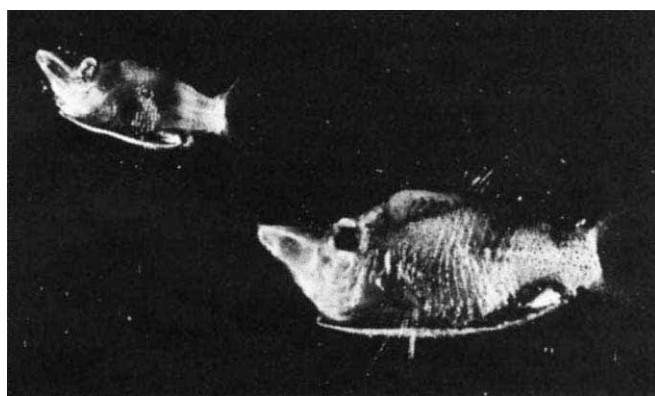


Fig. 5 Two specimens of the mesopelagic fish *Opisthoproctus grimaldii*. Luminous bacteria are maintained in a rectal diverticulum.

of diurnal changes in ambient light. Experimental evidence is now accumulating which indicates that in some animals at least the intensity of ventral bioluminescence can be varied. Free-swimming squid^{71,91}, ponyfish⁷² and luminescent fishes^{73,74} have been observed to alter their ventral luminescent output in response to changes in overhead light.

By what means does the animal ensure that its bioluminescent output matches the ambient light intensity? Many stomiatoid fishes have an orbital photophore that shines into the eye. It has been suggested that these organs allow the eye to 'light adapt' so that the fish is not blinded by its own emission⁷⁵. It seems more likely, however, that the function of this organ is to act as a reference standard for comparison with the light from above⁷⁶. Certainly, in hatchetfishes at least, the intensity of the orbital photophore waxes and wanes in concert with that of the ventral organs⁷⁷, a necessary feature of such a hypothesis. The supraorbital photophore in lanternfishes such as *Tarletonbeania* may also serve this purpose⁷³, while extra-ocular photoreceptors, the photic vesicles, may be similarly utilised in squids⁷⁸. Reference systems cannot however be identified in many fishes and other animals with ventral photophores. There seems no means, for example, whereby the upwardly directed eye of *Opisthoproctus* can detect the downwardly directed luminescence of its sole. It is possible that in such cases a fixed intensity emission may be modified by the responses of overlying chromatophores, and these in turn respond directly or indirectly to ambient light intensities. The colour and pattern responses of many shallow water fishes have an analogous feedback control requirement to the ventral bioluminescence of deep-sea species.

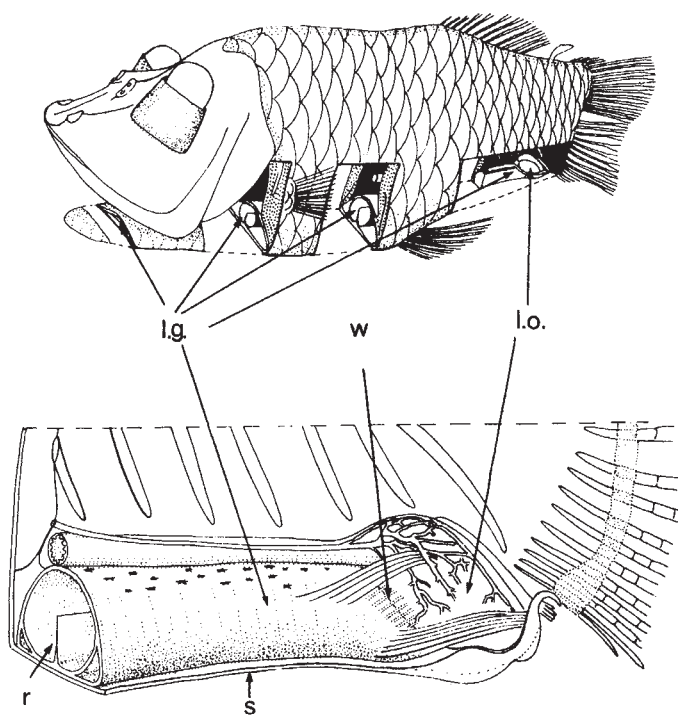
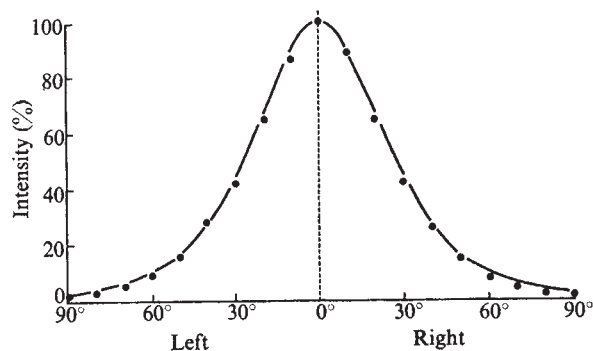


Fig. 6 The anatomy of the light organ of *Opisthoproctus soleatus* (from ref. 10). a, Lateral view showing the position of the light organ (l.o.) and light guide (l.g.). b, Detail of the light organ with its window (w) into the light guide, a hyaline tube surrounded by reflecting platelets (r) and diffusing the light over the whole area of the flattened sole (s).

Ventral counter-illumination can only be of camouflage value at depths in which downwelling light provides a significant background illumination, namely the epi- and mesopelagic zones⁶³. In the surface waters during the day the ambient intensity is too great to be countered by bioluminescence, and other aids to concealment, such as transparency, have to be used. Ventral illumination will be of greatest effect during the day in the mesopelagic zone, but at night will be equally effective much nearer the surface. The shallower living animals will need to emit more light than those in deeper and darker regions, and their respective ventral luminescent systems should give some indication of this. It was recognised very early that there is a correlation between the depth of a species and the development of its photophores⁷⁹. Murray and Hjort⁸⁰ noted that the fish *Cyclothone signata* had larger photophores than *C. microdon*, which was caught at greater depths. This is a

Fig. 7 A comparison between the angular distribution (in the transverse plane) of light from the hatchetfish *Argyrolepecus affinis* (●) and a computed curve for daylight in the sea (solid line) (from ref. 67).



general feature of many groups of animals, though it has only become clear with the advent of better techniques for the determination of the detailed vertical distribution of pelagic and planktonic organisms. Thus in *Cyclothone* the shallowest species *C. alba* and *C. braneri* have larger photophores than the deep *C. acclimens* and *C. microdon*. The deepest species, *C. obscura*, has only rudimentary photophores. Similar series are found in species of *Gonostoma*, *Stomias*, the hatchetfishes, and the lanternfishes. In the latter group the photophores of the deeper genera (for example, *Lampadena*) are much smaller than those of shallower genera (for example, *Hygophum*, *Diaphus*). Similar traits occur in other groups. The deepest species of the euphausiid genus *Thysanopoda* have relatively much smaller photophores than shallower species, and in the

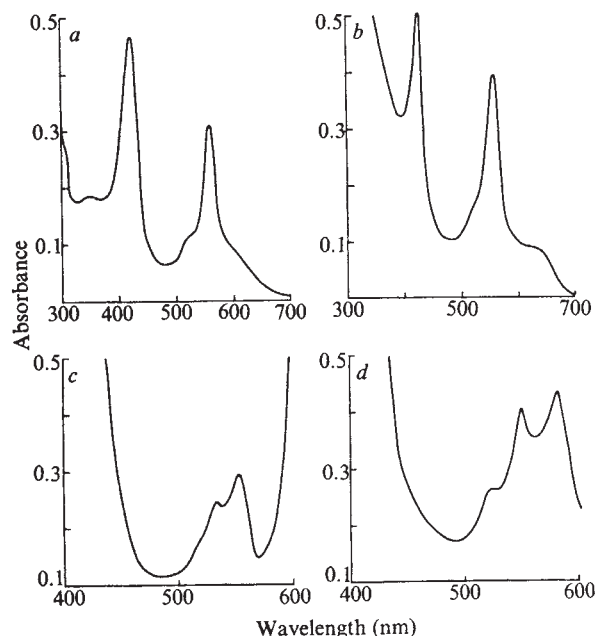


Fig. 8 Visible absorption spectra of pigments extracted from the photophore filters of four species of mesopelagic fishes. In each case the pigment is an effective blue filter, so that the final emission closely approximates the spectral distribution of ambient daylight in the sea. a, *Chauliodus sloani*; b, *Stomias boa*; c, *Valenciennellus tripunctulatus*; d, *Argyrolepecus aculeatus*.

shrimp genus *Systellaspis* the photophore size in different species also mirrors the depth distribution⁸¹. Some of the very deep living animals may use their photophores for camouflage only during their nocturnal migrations close to the surface.

Developmental information provides additional support. Many marine animals pass through a period as transparent larvae in the surface waters before sinking to greater depths as an adult. Those fish larvae that sink gradually during development tend also to develop their photophores gradually, with the result that they are appropriately equipped at all stages to mask their increasing opacity. Those that sink suddenly during a rapid metamorphosis tend to develop all their photophores at once, so that they, too, are equipped for their new light regime^{82,83}.

By no means all ventral photophores need be regarded as counter-illumination systems, and McAllister⁸⁴ has pointed out some of the problems inherent in any hypothesis of universal ventral camouflage. Many animals with ventral photophores are found close to the bottom during the day (for example ponyfishes). Counter-illumination would serve no purpose during the day—but might still be nearer the surface at night. Equally, if the orientation of of advantage in moonlit waters were the fish to migrate

a counter-illuminating animal changes in the water its camouflaging light could become a disastrous beacon. One of the stumbling blocks of a camouflage hypothesis is that submersible observations sometimes describe how many fishes seem almost randomly orientated in the water^{83,86}, though the animal may not necessarily be luminescing⁸⁷. Certain decapod and euphausiid shrimps can alter the orientation of their photophores^{88,89}, and this would be an effective means of compensating for changes in their orientation in the water. A further difficulty raised by bathysphere observations is the visibility of luminescing fishes. If their camouflage is effective they should not be visible. Yet Beebe⁹⁰ could "plainly see the lights of their (hatchetfishes) downwardly-directed photophores when the fish swam above me". These observations seem somewhat at variance with the camouflage concept of ventral luminescence. Nevertheless, there is now very considerable evidence, albeit still largely circumstantial, providing support for the hypothesis.

Observations of bioluminescence are always fascinating and often perplexing; their interpretations are even more so. Both observations and interpretations have a long history of suspicion, or outright disbelief. In the early eighteenth century a letter was received by the French Academy of Sciences from the people of Cadiz¹ "importing that . . . they had seen the whole sea shining with a clear light almost like a liquid phosphorus, . . . the seawater being carried away in bottles gave the same light in the dark; that some drops of it being let fall on the ground shone like sparks of fire, and that linnen dipt in this water became luminous." The Academy responded to this quite accurate description of dinoflagellate luminescence by retorting "The fact having been well examined, is found to be false, . . . The Academy think, they do as much service to the public, in disabusing them with regard to false wonders as in recounting to them the true." In which category would they have included the Puget Sound *Porichthys*, fishes with a complete battery of mainly ventral photophores—none of which work?

I thank Mr J. R. Badcock and Dr F. McCapra for commenting on the manuscript and Mrs C. Darter for preparing Figs 2 and 6.

- 1 Harvey, E. N. *A History of Bioluminescence* (American Philosophical Society Philadelphia, 1957).
- 2 Tizard, T. H., Moseley, H. N., Buchanan, J. Y. & Murray, J., *Rep. Scient. Results Voy. Challenger Narrative* 2, 743 (1885).
- 3 Hastings, J. W. & Wilson, T. *Photochem. Photobiol.* 23, 461–473 (1976).
- 4 Nealson, K. H. *Arch. Microbiol.* 112, 73–79 (1977).
- 5 Reichelt, J. L. & Baumann, P. *Arch. Mikrobiol.* 94, 283–330 (1973).
- 6 Nealson, K. H. & Hastings, J. W. *Arch. Microbiol.* 112, 9–16 (1977).
- 7 Ruby, E. G. & Nealson, K. H. *Biol. Bull. mar. biol. Lab. Woods Hole* 151, 574–586 (1976).
- 8 Tett, P. B. & Kelly, M. G. *Oceanogr. mar. Biol. A. Rev.* 11, 89–173 (1973).
- 9 Marshall, N. B. & Iwamoto, T. *Mem. Sears. Fdn mar. Res.* 1, 496–665 (1973).
- 10 Bertelsen, E. & Munk, O. *Dana Rep.* 62, 1–17 (1964).
- 11 Herrington, P. J. in *Proc. 9th Eur. mar. Biol. Symp.* (ed. Barnes, H.) 563–572 (Aberdeen University Press, Aberdeen, 1975).
- 12 O'Day, W. T. *Contr. Sci.* 255, 1–12 (1974).
- 13 Bertelsen, E. *Dana Rep.* 39, 1–256 (1951).
- 14 Meyer-Rochow, V. B. *Experientia* 32, 1175–1176 (1976).
- 15 Boufield, E. L. & Klawe, W. L. *Bull. Sth. Calif. Acad. Sci.* 62, 1–8 (1963).
- 16 Hastings, J. W. *Oceanus* 19, 17–27 (1976).
- 17 Haneda, Y. *Pacif. Sci.* 5, 372–378 (1951).
- 18 Rosenblatt, R. H. & Montgomery, W. L. *Copeia* 510–515 (1976).
- 19 Haneda, Y. & Tsuji, F. I. *J. Morphol.* 150, 539–552 (1976).
- 20 Morin, J. G. *et al. Science* 190, 74–76 (1975).
- 21 Dubois, R. *C. r. heb. Séanc. Acad. Sci., Paris* 105, 690–692 (1887).
- 22 Harvey, E. N. *Bioluminescence* (Academic, New York, 1952).
- 23 McCapra, F. *Endavour* 32, 139–145 (1973).
- 24 Johnson, F. H. & Shimomura, O. *Bioscience* 25, 718–722 (1975); in *Photophysiology* (ed. Giese, A. C.) 8, 275–334 (Academic, New York, 1972).
- 25 Shimomura, O., Johnson, F. H. & Saiga, Y. *J. cell. Physiol.* 59, 223–239 (1962).
- 26 Goto, T., Iio, H., Inoue, S. & Kakoi, H. *Tetrahedron Lett.* 2321–2324 (1974).
- 27 Shimomura, O. & Johnson, F. H. in *Symp. Soc. exp. Biol.* 30, 41–54 (1976).
- 28 Shimomura, O. & Johnson, F. H. *Science* 159, 1239–1240 (1968).
- 29 Hastings, J. W. & Morin, J. G. *Biochem. biophys. Res. Commun.* 37, 493–498 (1969).
- 30 Cormier, M. J., Hori, K. & Anderson, J. M. *Biochim. biophys. Acta* 346, 137–164 (1974).
- 31 Shimomura, O. & Johnson, F. H. *Biochemistry* 11, 1602–1608 (1972).
- 32 Hori, K., Cormier, M. J. *Proc. natn. Acad. Sci. U.S.A.* 70, 120–123 (1973).
- 33 Cormier, M. J. *et al. J. cell. Physiol.* 81, 291–297 (1973).
- 34 Hastings, J. W. *A. Rev. Biochem.* 37, 597–630 (1968).
- 35 Shimomura, O. & Johnson, F. H. *Nature* 256, 236–238 (1975).
- 36 Tsuji, F. I., DeLuca, M., Boyer, P. D., Endo, S. & Akutagawa, M. *Biochem. biophys. Res. Commun.* 74, 606–613 (1977).
- 37 Kishi, Y., Goto, T., Hirata, Y., Shimomura, O. & Johnson, F. H. *Tetrahedron Lett.* 3427 (1966).
- 38 Inoue, S., Kakoi, H. & Goto, T. *Tetrahedron Lett.* 2971–2974 (1976); *Chem. Commun.* 1056–1057 (1976).
- 39 Yamaguchi, I. *Biochem. J.* 151, 9–16 (1975).
- 40 McCapra, F. & Manning, M. J. *Chem. Commun.* 468 (1973).

- 41 Johnson, F. H., Sugiyama, N., Shimomura, O., Saiga, Y. & Haneda, Y. *Proc. natn. Acad. Sci. U.S.A.* 47, 486–489 (1961).
- 42 Sie, E. H. C., McElroy, W. D., Johnson, F. H. & Haneda, Y. *Archs Biochem. Biophys.* 93, 286–291 (1961).
- 43 Cormier, M. J., Crane, J. M. & Nakano, Y. *Biochem. biophys. Res. Commun.* 29, 747–752 (1967).
- 44 Tsuji, F. I. & Haneda, Y. *Nature* 233, 623–624 (1971).
- 45 Haneda, Y., Johnson, F. H. & Shimomura, O. in *Bioluminescence in Progress* (eds Johnson, F. H. & Haneda, Y.) 533–545 (Princeton University Press, Princeton, 1966).
- 46 Tsuji, F. I., Haneda, Y., Lynch, R. V. & Sugiyama, N. *Comp. Biochem. Physiol.* 40A, 163–179 (1971).
- 47 Tsuji, F. I., Barnes, A. T. & Case, J. F. *Nature* 237, 515–516 (1972).
- 48 Barnes, A. T., Case, J. F. & Tsuji, F. I. *Comp. Biochem. Physiol.* 46A, 709–723 (1973).
- 49 Tsuji, F. I. *et al. Molec. cell. Biochem.* 9, 3–8 (1975).
- 50 White, E. H., McCapra, F., Field, G. F. & McElroy, W. D. *J. Am. chem. Soc.* 83, 2402 (1961).
- 51 Shimomura, O., Johnson, F. H. & Kohama, Y. *Proc. natn. Acad. Sci. U.S.A.* 69, 2086–2089 (1972).
- 52 Ohtsuka, H., Rudie, N. G. & Wampler, J. E. *Biochemistry* 15, 1001–1004 (1976).
- 53 Herring, P. J. *Proc. R. Soc. Edin.* 73B, 229–238 (1972).
- 54 Seliger, H. H. & McElroy, W. D. *Proc. natn. Acad. Sci. U.S.A.* 52, 75–81 (1964).
- 55 Harvey, E. N. *Am. J. Physiol.* 70, 619–623 (1924).
- 56 Morin, J. G. & Hastings, J. W. *J. cell. Physiol.* 77, 313–318 (1971).
- 57 Wampler, J. E., Hori, K., Lee, J. W. & Cormier, M. J. *Biochemistry* 10, 2903–2909 (1971).
- 58 Wampler, J. E., Karkhanis, Y. D., Morin, J. G. & Cormier, M. J. *Biochim. biophys. Acta* 314, 104–109 (1973).
- 59 Dahlgren, U. *J. Franklin Inst.* 181, 525–556 (1916).
- 60 Rauther, M. in *Bronn's Klassen und Ordnungen des Tierreichs* 6, 125–167 (1927).
- 61 Fraser, J. *Nature Adrift* (Foulis, London, 1962).
- 62 Clarke, W. D. *Nature* 198, 1244–1246 (1963).
- 63 Clarke, G. L. & Denton, E. J. in *The Sea* 1, (ed. Hill, M. N.) 456–468 (Wiley-Interscience, New York, 1966).
- 64 Haneda, Y. *Rec. oceanogr. Wks Japan* 7, 83–87 (1963).
- 65 Denton, E. J., Gilpin-Brown, J. B. & Roberts, B. L. *J. Physiol., Lond.* 204, 38 (1969).
- 66 Denton, E. J. *Phil. Trans. R. Soc.* 258B, 285–313 (1970).
- 67 Denton, E. J., Gilpin-Brown, J. B. & Wright, P. G. *Proc. R. Soc.* 182, 145–158 (1972).
- 68 Denton, E. J., Gilpin-Brown, J. B. & Wright, P. G. *J. Physiol., Lond.* 208, 72–73 (1970).
- 69 Boden, B. P. & Kampa, E. M. *Symp. zool. Soc. Lond.* 19, 15–26 (1967).
- 70 Clarke, T. A. *Fish. Bull. U.S.* 71, 401–434 (1973).
- 71 Young, R. E. & Roper, C. F. E. *Science* 191, 1046–1048 (1976).
- 72 Hastings, J. W. *Science* 173, 1016–1017 (1971).
- 73 Lawry, J. V. *Nature* 247, 155–157 (1974).
- 74 Case, J. F., Warner, J., Barnes, A. T. & Lowenstein, M. *Nature* 265, 179–181 (1977).
- 75 Marshall, N. B. *Aspects of Deep Sea Biology* (Hutchinson, London 1954).
- 76 Nicol, J. A. C. *Symp. zool. Soc. Lond.* 19, 27–55 (1967).
- 77 Herring, P. J. in *A Voyage of Discovery* (ed. Angel, M. V.) (Pergamon, Oxford, in the press).
- 78 Young, R. E. *Deep-Sea Res.* 19, 651–660 (1972); *Pacif. Sci.* 27, 1–7 (1973).
- 79 Brauer, A. *Wiss. ergebn. dt. Tiefsee-Exped. 'Valdivia'* 15, 1–175 (1908).
- 80 Murray, J. & Hjort, J., *The Depths of the Ocean* (MacMillan, London, 1912).
- 81 Herring, P. J. *J. mar. biol. Ass. U.K.* 56, 1029–1047 (1976).
- 82 Ahlstrom, E. H. in *The Early Life of Fishes* (ed. Blaxter, J. H. S.) 659–674 (Springer, New York, 1974).
- 83 Badcock, J. R. in *A Voyage of Discovery* (ed. Angel, M. V.) (Pergamon, Oxford, in the press).
- 84 McAllister, D. E. *J. Fish. Res. Bd. Can.* 24, 537–554 (1967).
- 85 Barham, E. G. in *Proc. Int. Symp. biol. Sound Scattering in the Ocean* (ed. Farquhar, G. B.) 101–119 (Maury Center for Ocean Science, Washington, D.C., 1971).
- 86 Bone, Q. J. *J. mar. biol. Ass. U.K.* 53, 619–633 (1973).
- 87 Backus, R. H. *et al. Science* 160, 991–993 (1968).
- 88 Omori, M. *Adv. mar. Biol.* 12, 233–324 (1974).
- 89 Hardy, M. G. *Nature* 196, 790–791 (1962).
- 90 Beebe, W. M. *Half Mile Down* (Bodley Head, London, 1935).
- 91 Young, R. E. & Roper, C. F. E. *Fish Bull. U.S.* 75 (in the press).



A hundred years ago

RAINFALL OBSERVATIONS IN THE EAST OF FRANCE FROM 1763 TO 1870.—In the *Bulletin Hebdomadaire* of the Scientific Association of France, of the 10th instant, Prof. Raulin gives an interesting historical account of all the rainfall observations made during these 108 years anywhere in that section of France which is marked off by lines joining Givet on the Meuse, Lauterbourg on the Rhine, Belley near the Rhon, and Decize on the Loire, and which thus comprehends seven well-marked regions, viz., the plain of Alsace, the chain of the Vosges, the plateaux of Lorraine and Bourgogne, the plains of Champagne and Bresse, and finally the chain of the Jura mountains. During the past three years Prof. Raulin has been engaged collecting all available materials for a monograph on the rainfall of this part of Europe, which, judging from his great monographs of the rainfall of other sections of France and of the rainfall of Algeria, will doubtless take its place as a permanent contribution of very high value to meteorological science.

Marine research in China

Bertil Öström

National Board of Fisheries, Hydrographic Department, Fack, 40310 Göteborg, Sweden

This paper is a personal travel report from a visit to the People's Republic of China in September and October 1976. Oceanographic research, as seen from visits to institutes, participation in fieldwork and personal contacts, is discussed, with some emphasis on the physical and chemical aspects. There is a tangible striving for utility in research policy. Research activities are also to a large extent integrated with other functions of society. Marine research is of increasing importance in China today; there is a growing cognisance of the hazards involved in pollution of the marine environment. The technical standard varies; some institutes are equipped with older apparatus based on manual methods, whereas others had access to modern analytical instruments. Information about the important centres for oceanographic research, at Tsingtao on the Yellow Sea coast and at Kwangchow for the south China Sea is presented.

The sea is of great historical importance in Chinese society. It has been a temptation as an escape out of a land locked by mountains to the west. It has also carried enemies in warships to the long China coast.

The Chinese have long been eminent seafarers. The most renowned of them was Cheng Ho, alias San Pao (he is best known under the latter name in Asia and on the Southern islands). From the mouth of Yangtze Kiang he started enormous trade expeditions with 500-foot ships and more than 20,000 men in the beginning of the fifteenth century, that is before Diaz, Vasco da Gama, Columbus and Magellan. Cheng reached India, Persia and Ethiopia. On later voyages he visited Zanzibar to the East of Africa and Timor near Australia.

Later there was a decline in Chinese seafaring. Internal antagonism during the Ming dynasty led to a prohibition of long voyages and all ships with more than two masts were broken up. The Chinese command of the sea was thereby irretrievably lost. The Portuguese and the Dutch soon became the new commanders of the seas. Later there followed a period of lively sea traffic to China (but without Chinese participation), and during the following centuries parts of the Chinese coast were colonised. Not until this century with the liquidation of most of the colonies (some of them, for example Hong Kong, are still left) has a new Chinese merchant navy grown up. Today this merchant navy is of a considerable size, often sailing under the flag of convenience of Somalia or Panama. Local river navigation has always been important and the transport of goods by junks has been the core of the Chinese transportation network.

The sea has long been a source of nutriment for the Chinese. The profession of the fisherman is mentioned in the oldest documents. Even more characteristic of the old Chinese knowledge of the sea is the use of marine algae for food. Algae are collected at sea and included in the fare where they are either eaten directly or cooked. Some

are preserved or dried and stored for later use. More than a hundred species are used for food.

Supply of resources from the sea

Interest in the sea is today focused on natural resources. The question, what can the sea give up in the way of food, minerals and energy? is being raised all over the world. When it comes to extraction of food from the sea we are today departing from the hunting stage where we catch whatever we can, to the farming stage where we plan and breed aiming at maximal yield but with protective measures taken to support the regrowth. In this field China has a great advantage in the knowledge that is accumulated from freshwater farming. For many thousands of years the Chinese people have collected information about fish breeding in ponds, of how to master the diseases of fish and of how to best feed and catch. Some of this knowledge can be transferred to conditions in the sea. Regarding

Fig. 1 Sketch map of China. The coastline is approximately 10,000 km long. Tsingtao on the Shangtung peninsula is a centre for marine research. Large mussel farms are kept in the shallow water of the Yellow Sea outside the town and marine algae are cultivated on beds in the tidal zone close to the shore. The Pohai Sea and Yellow Sea are steadily silting up (Shangtung was formerly an island and Peking was situated at the sea) due to mud deposits from the northern rivers. Kwangchow has several marine institutes and is now a centre for present research activities in the South China Sea.



minerals and energy, China, like many other countries, is now aware of the potential richness of the sea. Sea areas which were previously fairly uninteresting become suddenly valuable and territorial demands are raised and fishing rights are claimed.

Marine research

In these circumstances it is natural that marine research in China is expanding. Research institutes are growing, new vessels are built, field stations are established and new problem areas are investigated. My journey to China with visits to various marine institutes was arranged to study marine research and to get a view of its status today. It was a research exchange between China and Sweden and a group of Chinese scientists had the opportunity to study the corresponding Swedish conditions.

A comparison of research policies shows difference of principles between China and the 'Western countries'. In the West we have an extensive basic research and a parallel applied research which sometimes, when needed, uses results from the basic research. This basic research is sometimes termed free research and the applied research is called



Fig. 2 The research vessel, The Red Star, anchored on the Pearl River near Kwangchow (Canton): 66 m length, 11 m width, 3.7 m depth, 1500 tons dead weight, 12 knots. It takes 100 tons of freshwater and 120 tons of fuel allowing expeditions of up to 3 weeks at sea with a crew of 35 (sailors and scientists). Many of the cruises are directed to the South China Sea. Home harbour is Kwangchow.

technical research. Free research is based on personal interests of individual scientists who find a certain problem area interesting. These scientists apply for finance from the scientific boards directing the research, are supplied with grants and make an investigation resulting in a report. In this way knowledge of previously unknown conditions is documented to form a field of knowledge where the gaps, one after the other, are filled in. The individual scientist at the time of the investigation may not necessarily see of what direct use the knowledge he accumulates can be. It is known empirically that basic research is profitable by creating a base for technical development, and the scientist can assume that the knowledge might come to use some time in the future.

This so called free research is not emphasised in China. Instead, research is strongly directed at problems. The investigation is expected to give a practical result as soon as possible and is therefore more like our technical research. The reason for this probably lies in Chinese modern history,

where the overall dominating aim was to free the population of over 800 million from starvation (a task carried out within one human generation). In this situation all marine research was aimed at an increase in the yield of nutriment from the sea, and it is possible that increased basic research will return with increased economic margins. It is probable that a weak basic research in the long run will check the applied research. On the other hand, the question can be raised as to whether we in the Western countries sometimes solve hypothetical problems that consume resources but are of no practical importance. Probably we have a lot to learn from each other. If we learn and take advantage of the pronounced utility in the Chinese research and the broad base for technology provided by the Western research, we should find a good formula for truly effective research.

A distinct impression from the collaboration with the Chinese scientists is that they are well aware of the events and progress of science in the West, but that they carefully sort out and make use of only the results that in some way are fit and good for China. Many of the senior scientists have, after the liberation, returned from institutions in USA and Europe. They master foreign languages, have a good overview of the research situation in the Western countries and today they often have leading positions in their institutes. The flow of information from Western literature is sufficient and the institution libraries often have the later important works in the field of marine research. Yet there are no signs of Americanisation or strong European influence in research policy and methods. As a visitor one gets, after a while, a strange feeling of being 'looked through'; one realises that the Chinese colleagues fairly well know one's situation without being interested in taking it over.



Fig. 3 Shanghai harbour. To the right the fore-end of a modern trawler. The signs mean Shanghai Fishing Company and are followed by a number. In the background modern trade ships and in front of them traditional wooden sailing junks carrying goods up the Yangtze Kiang river.

There is also an interesting difference regarding research means and methods. We can, in our countries, hardly purge ourselves from new ideas and methods which are pushing on, regardless of if we really should prefer them or not, whereas the Chinese society has found ways completely to control and direct outside influence.

There is a great interest in China throughout the world. Academia Sinica who mediate the scientific visits receive many applications from groups wanting to visit China, and they strain off a considerable part. But there is today a great Chinese interest in Western methods, knowledge and experience in the areas of pollution studies and environmental protection.

Much of the equipment at the research institutes consists of simple, sometimes almost primitive aids, much apparatus is worn out and also the basic resources like buildings and fixtures are often old and worn. On the whole, the technical standard is quite low. This depends partly on an actual low standard, but partly also on the thinking based on optimisation and utility that is found everywhere in China; human and material resources should be used in a judicious way, overdone technology has no mission; all activities should aim at an improvement of the vital conditions for the people.



Fig. 4 Buoy systems carrying nets on which blue mussels grow in the mussel farms south of Tsingtao in the Yellow Sea. (Photograph by Hans Ackefors).

On the other hand there is no lack of resources 'when they are needed', that is, when a problem is considered urgent. If, say, an unknown disease were to threaten the huge marine cultivation farms on the coast, large resources of personnel, research vessels and modern instruments would be available to attack the problem. Laboratory instruments were sometimes amazingly modern in design and function. As an example there were Chinese manufactured scanning recording spectrophotometers and gas chromatographs with linear temperature programming and all types of interchangeable detectors including electron capture. The instruments had also advanced industrial design in their performance, which is an indication of a high technical level. These instruments were manufactured in Shanghai.

Chinese marine research has access to a number of research vessels making regular cruises at sea. The Yellow Sea, being a productive shallow water area on the shelf has traditionally seen the main part of the research activities with a centre in Tsingtao on the Shangtung peninsula. The East China Sea coast has some importance for fishing and cultivation of algae. Today, however, the interest is largely concentrated on the South China Sea. This is caused by the increased awareness of the fish resources in the area and the recent discovery of potential oil and mineral resources. Field stations have been established on previously unoccupied groups of islands and an intensive investigation programme concerned with the marine geology, hydrography and fishery biology of the area has started. Unfortunately, it was not possible to visit the most interesting research vessel, the Red Star, as the ship was on a cruise to the South China Sea. The photograph, Fig. 2, was taken by another Swedish visitor (Olof Tandberg); it shows the Red Star which is approximately 2,000 tons dead weight.

Information was also obtained on two newly built ocean-going research vessels of approximately 10,000 tons dead weight, which were at that time on expeditions in the Pacific. These ships are controlled directly by the central government in Peking. Information on these ships has also been given in the information journal *Peking Review*, published in English in several countries. The ships have equipment for marine geological investigations and laboratories for chemical analyses. There are also many smaller research vessels, many of which operate in connection with field stations on the coast. In addition, research work and practical work are to a large extent integrated in China, so that, for example, fishing boats take samples for later analysis at the research laboratories.

Visits to marine research institutes

Practical oceanographic research work was studied during visits to a number of research institutes along the coast. A certain formalism and conformity characterised all such visits. The visit is introduced with a quiet chat where everybody is drinking tea, sitting in a rectangular formation along three of the four walls of the room. The talk is, however, mainly directed by the group leader of the visitors and the head of the reception committee, and does not have the form of a general discussion.

During continued tea drinking the political objectives of the institute are then thoroughly described by the political representative of the reception committee. It is obvious that chairman Mao's thinking has had and still has an immense influence on all activities in China. All institutes visited lean back on his statements as directives for their work. Some political concepts often recur. One such concept is the 'trinity' which can refer to peasants, workers and soldiers, for example, in discussions of recruiting students to the universities, or it can sometimes denote other constellations such as workers, technicians and supervisors in discussions of a technical problem. The fact that an instrument is constructed by an engineer, the assembly is planned by a supervisor and performed by a worker is in our language convention understood; we know it, but we do not say it. In China, it is felt important to underline such facts, to mark a new approach, a difference from the past, and when displaying an instrument these three anonymous but equal originators are mentioned.

Another concept is 'open-door research' which underlines the cooperation between research and production. The scientist should not retire into splendid isolation when solving his problems but carry out his research in the industries cooperating with the workers and considering their practical viewpoints of the problem.

After the information on the political objectives follows an account of the scientific research program, given by a scientist of the institute. Thereafter follows a guided visit to the different laboratories to see the activities and equipment. There is an open atmosphere during these visits, questions are answered, everything is shown on request, improvisations and suggested changes of the program are acceded to as far as is reasonable. (This was also the case for the free time during the programme. There were no restrictions on the movements and activities of the visitors.)

Apparatus and technical equipment

Chinese research laboratories have different, not so strict, partitioning into disciplines compared with that which we are used to. Problems which we would have split up into, say, a biochemical and a marine geological part, may in China be solved at the same place. The partitioning of fields of activities and translations of laboratory names made in the following are therefore often approximate and difficult to bring close to reality.

The research vessel the Red Star (Fig. 2) has a common water sampling equipment, with deep-going winches and water samplers, instruments for observation and measurements of physical properties, tube and gram samplers for sediments and laboratory on board for chemical analysis. Part of the equipment is carried on board for every expedition, for example, the photometer. There is also a laboratory for treatment and conservation of biological samples.

The institutes visited all had laboratories for analysis of water samples which were brought back from sea. Shantung Institute for Oceanography had a modern salinometer, based on electric conductivity of the sea water, of high sensitivity ($S = \pm 0.001\%$). This institute manufactured for calibration purposes, standard sea water which was sealed into glass ampoules.

As mentioned above, some institutes were equipped with modern gas chromatographs and spectrophotometers. Access to mass spectrographic analysis was also provided, although we did not see an installation during the visits. At Shantung Institute for Aquatic Products a cold atomic absorption instrument for a.o. mercury determinations was demonstrated.

The sampling equipment shown at some institutes was of conventional types. Russian designed grab samplers for sediment were shown in Kwangchow, and common types of plankton nets were also to be seen.

The basic equipment of the laboratories such as centrifuges, drying ovens, microscopes, water thermostat baths and oscilloscopes was in the main similar to what we are used to in our laboratories for routine oceanographic analyses.

The carbon-14 technique for measurement of the primary phytoplankton production was known and was studied at Hupeh Hydrobiological Institute at Yangtze Kiang. It was not yet in use at sea, however, and the primary production was determined according to the oxygen method using light and dark bottles.

Remote sensing as a tool in marine research is as yet very little used in China. The scientific element of the country's own satellite program is (according to verbal communication) merely used for studying meteorological problems. It has been said that China is probably taking information from the LANDSAT-satellites, but comments on that were not received.

Pollution problems are receiving an increasing interest as a result of the increased pollution by industrialisation. The scientists are aware of the problems and anxious to learn from other countries and share their experiences of solutions to the problems. It is unlikely that they will allow a mercury and cadmium catastrophe like that in Japan.

It is admitted, however, that DDT is a serious problem and that even deaths have occurred. Intensive research in the area of biological control is going on in order to replace the chemical biocides with biological control methods. Instrumentation for routine analysis of heavy metals and for DDT was seen at many laboratories.

The Institutes

The Institutes carrying out oceanographic research were the following.

South China Sea Institute of Oceanography in Kwangchow (Canton), is administered by Academia Sinica. It specialises in the South China Sea and has laboratories specialising in hydrography and meteorology, marine geology, marine biology and hydrophysics.

The institute was founded at Chanchiang on the south coast but was moved to Kwangchow in 1973. This institute

is administering the research vessel, the Red Star. It has also been engaged in pollution studies. There are 340 employees, including those in field stations. A new building for the institute comprising six floors of a total of 9,000 m² for research and administration is under construction. Another research vessel of about 1,000 tons is being built this year.

Kwantung Institute for Aquatic Products in Kwangchow. The institute specialises on the South China Sea. The main task is to develop and utilise the resources of the South China Sea. It has about 500 employees, most of them stationed on field stations belonging to the institute. A new building is under construction for the mother institute in Kwangchow, which will contain a political office, an administrative office and six laboratories. The institute has several smaller research vessels. One branch of their activities is also to spread scientific information among the people.

Shanghai Institute for Aquatic Products in Shanghai. The Institute had, after the cultural revolution 1968, no research vessels of its own, but its scientists are working from fishing boats in cooperation with the fishermen and the fishing company (open-door research). About 200 are employed. The Institute has a laboratory for marine resources.

Tsingtao Institute of Oceanology in Tsingtao. The institute belongs to Academia Sinica. Their working area is the entire coast from Pohai to the South China Sea. About 500 persons are employed. They have three research vessels of 1,000, 400 and 300 tons dw respectively and laboratories for marine physics, marine geology, marine chemistry and oceanographic research instruments. The institute is also engaged in pollution studies. Research is to some extent concentrated on the Yellow Sea shelf area. Sediment movements are also studied.

Shantung Institute of Oceanography in Tsingtao. The institute is mainly an educational institution. There are departments of marine hydrology and meteorology, marine physics, marine chemistry, marine geology and marine biology. There are about 500 employees and the number of students varies during the year. There are 840 places for a 3-yr education for marine scientists. There are also about 1,000 students on short term courses. Education is also given by correspondence. The institute has stressed work on pollution and environmental problems. For field work and excursions there are a number of small motor-boats.

Shantung Provincial Institute for Marine Products in Tsingtao, with about 220 employees. Working area is the Pohai Sea and Yellow Sea. Emphasis is on the living resources. Research is on effects of industrial sewage water on the marine environment.

Direct reference to Chinese scientists, and further details of marine research in China may be obtained from the author on request.

I thank The Royal Swedish Academy of Sciences and the Swedish Academy of Engineering Sciences for a fellowship, and Academia Sinica in Peking, for a well planned programme. Thanks are also due to the tour conductor, Mr Lu Ching-ling, and the interpreter, Mr Chu Chian-ning, and to Drs Bern Dybern, Hans Ackefors and Anders Jarmen and to Ann-Louise Martin for many fruitful discussions. Dr Olof Tandberg kindly submitted photographic and scientific material, and Mr Bertil Kuhlemann provided much useful information.

articles

Possible relation between optical and very long baseline interferometer radio observations of the QSO 3C 345

T. D. Kinman

Kitt Peak National Observatory, Tucson, Arizona 85726

Existing radio and optical data on the variability of the QSO 3C 345 are summarised. Interesting correlations between the polarisation in radio and optical bands are evident; moreover successive outbursts have similar polarisation. This suggests that all outbursts are triggered by activity in the same central object rather than being a consequence of unrelated explosions of supernovae or similar objects.

SEVERAL QSOs show an apparent superrelativistic expansion in the radio which often continues for several years along the same position angle. On one interpretation, the expansion must be guided by a relatively permanent structure, and I have, therefore, looked for other signs of this structure. During outbursts, the annual mean plane of polarisation of the optical radiation from 3C 345 aligns with the direction of the superrelativistic expansion: the magnetic field associated with the optical radiation is therefore perpendicular to the expansion. There is weaker evidence that in the radio the average intrinsic plane of polarisation is perpendicular to the expansion vector. These regularities give more support to the idea that a QSO has an ordered and relatively stable structure than that it is made up of an array of unrelated sources.

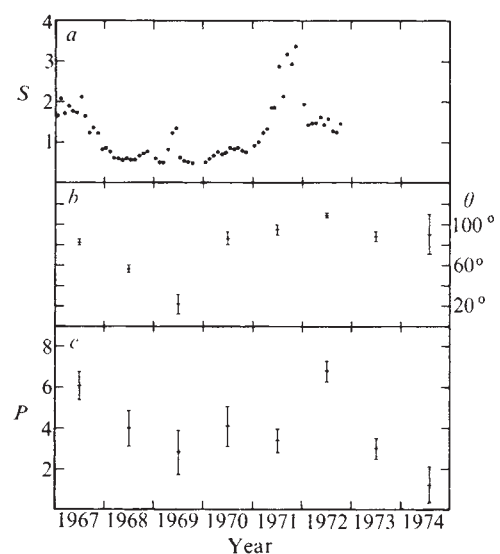
Superrelativistic expansion

3C 345 (redshift 0.595) is one of the three QSOs in which radio variability was first discovered by Dent¹. Its optical variability² was discovered soon afterwards, and also its associated property of a variable optical polarisation³. It has therefore been one of the most closely monitored sources at both optical and radio wavelengths. Recently, very long baseline interferometer (VLBI) radio observations at 2.8 cm by Cohen *et al.*⁴ and by Wittels *et al.*^{5,6} at 3.8 cm have shown that apparent superrelativistic expansion occurs in this source. Between 1974.15 and 1976.13, Cohen *et al.* found a separation increase of from 1.23 to 1.70 milliarc s on a two-component model. Wittels *et al.* found a separation increase of from 1.0 to 1.3 milliarc s from late 1971 to 1974 at the longer wavelength. In both cases, the position angle of the expansion was well defined with a position angle $105^\circ \pm 5^\circ$. Explanations of such expansion have been suggested in terms of (1) relativistic motion along the line of sight⁷, (2) a static array of sources whose flux variations simulate the expansion⁸, or (3) a model in which it is the phase motion of an actual expansion that

is observed^{9,10}. The second explanation requires that contractions should be observed as frequently as expansions, and this is not the case as far as is known. Sanders¹¹ has described a model of the third type in which the superrelativistic effect is produced by synchrotron emission from electrons travelling in a dipole magnetic field as seen from its equatorial plane. Each expansion originates from a relatively impulsive injection of relativistic electrons from a source near the centre of the dipole. The energy of these electrons is less than the energy of the magnetic field, and so the field survives, and the same expansion orientation may be expected for successive outbursts.

It seems likely that the impulsive injection of the electrons should be associated with an increase in the optical flux. A simple linear extrapolation back to zero component separation suggests that the present expansion originated between 1958 and 1963 (3.8-cm observations) or at about 1967 (2.8-cm observations). The time of origin is therefore quite ill-determined. Furthermore, if the beaming angle of the optical flux is very small (1–10 mrad) as suggested by Cocke and Pacholczyk¹², the actual optical emission associated with a specific radio expansion might

Fig. 1 Optical flux and linear polarisation properties of 3C 345: *a*, monthly means of the optical flux in mJy; *b*, annual means of the position angle θ of the polarised flux; and *c*, annual mean of the percentage polarisation.



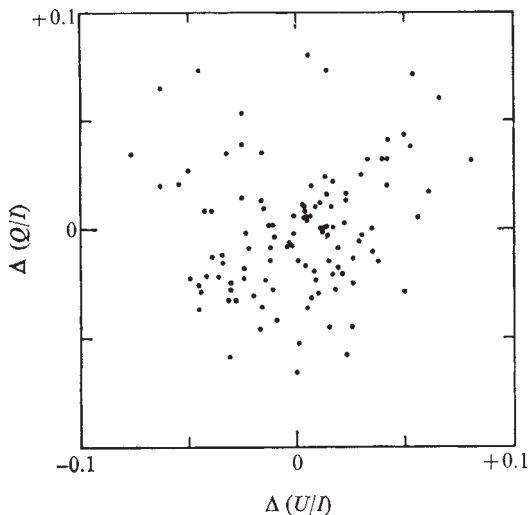


Fig. 2 The residuals of the normalised Stokes parameters for the optical flux from 3C 345 after removal of their annual mean values.

be missed if it did not occur in the direction of the Earth. But even in this case, if the model has a permanent magnetic structure, each optical event should carry in its linear polarisation the signature of the magnetic field geometry and this should be related to the geometry of the expansion.

Optical linear polarisation

Some 114 observations of the optical linear polarisation of 3C 345 have been made between 1967 and 1974, inclusive, by Elvius¹³, Visvanathan¹⁴, Babadzhanyan *et al.*¹⁵, Kinman *et al.*¹⁶, and unpublished observations by the author. This polarisation is characterised by quite large changes which can occur on time scale as short as a few days. To smooth out such changes in a search for systematic effects, annual means of the normalised Stokes parameters (Q_m/I , U_m/I) were formed and used to compute the annual mean polarisation P and the annual mean position angle θ of the optical E-vector. These are shown in Fig. 1; the error bars are the r.m.s. deviations of these means. The distribution of the residuals ($\Delta[Q/I]$, $\Delta[U/I]$) from these means is shown in Fig. 2. Systematic trends (for example, skewness in the distribution of these residuals) may be present, but seem small enough for the means (Q_m/I , U_m/I) to have enough significance for our present purpose. These annual means and their residuals corre-

spond roughly to the slowly varying component B and the rapidly varying component C in the more elaborate analysis by Kinman *et al.*¹⁶. In the top panel of Fig. 1, the monthly means of the optical flux S are given in mJy; they were derived from B -magnitudes given by Kinman *et al.*¹⁶, Barbieri and Abati-Erculiani¹⁷, Lü¹⁸, Babadzhanyan *et al.*¹⁵, Tritton and Selmes¹⁹, and unpublished observations by the author using the relation

$$\log_{10} S = 6.67 - 0.4 B$$

Two substantial outbursts are seen. One began in October 1966 (Kinman *et al.*¹⁶) and ended in 1968, and the second has its maximum in the second half of 1971. The sparse data available from old plates (Angione²⁰) do not contradict the idea that outbursts of this kind are occurring at a rate of about one every five years. It has been suggested by C. Barbieri, G. Romano, S. di Serego and M. Zambon (personal communication) that an actual period of 4.38 yr is present. It does not seem unlikely that these outbursts are associated with the large injections of electrons that give rise to the expansion phenomenon on Sander's¹¹ model. An injection event rate much higher than this would presumably result in the superposition of several expanding radio sources at one time. If the rate was much lower, the expansion phenomenon would be rarer than we observe.

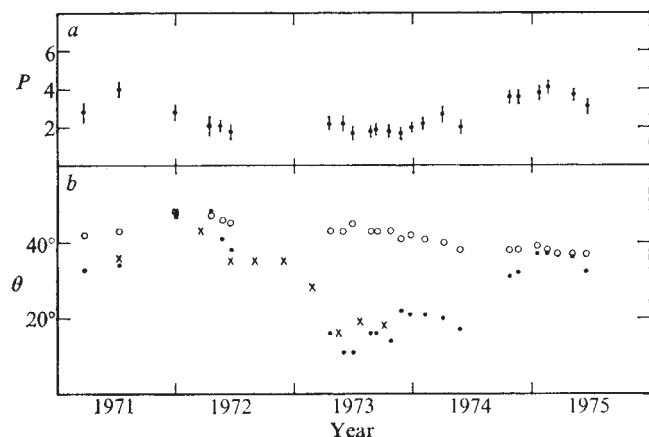
The optical position angle shows some regularity in its behaviour—lying between 80° and 110° when these major outbursts occur. The emission in a highly polarised source of synchrotron radiation must come from an optically thin plasma and will have the E-vector perpendicular to the magnetic field. The field associated with the optical radiation is therefore perpendicular to the expansion direction and (on Sander's model) to the axis of a dipole field. If we speculate that the central source is a massive body with its rotation axis parallel to the axis of the dipole field, the optical radiation could come from an equatorial field on the massive object. In order to account for periodic effects in the optical flux of 3C 345, Kinman *et al.*¹⁶ speculated that the slowly varying component B (which is large when the optical flux is high) arises in the equatorial zone of a rotating body.

Radio polarisation

Radio polarisation data from 1971 to 1975 are shown in Fig. 3. A rotation measure of 18.3 rad m^{-2} (Haves²¹) was applied to the 3.7- and 11.1-cm data of Altschuler and Wardle²², and the 6-cm data of Seielstad and Berge²³. On Sander's model, we would expect the relevant magnetic field to lie predominantly parallel to the expansion direction, and so would expect a position angle for the radio linear polarisation of 15° if the plasma is optically thin, or 105° in the optically thick case. If the field is a permanent structure, the position angle should not show a large secular trend, and this is observed. While the optically thin case does agree better with the observations, it does not seem possible to apply the usual criteria of an expanding cloud model (particularly if beaming effects are present) to decide whether the source is, in fact, optically thin or not. We only note that if it is optically thin for most of this time, then the excursion of the position angle for the 3.7- and 6.0-cm data during 1973 and 1974 (and the simultaneous drop in the percentage polarisation) can be understood in terms of the transient appearance of an optically thick component at these wavelengths.

The present optical and radio data are therefore compatible with models of 3C 345 which involve a relatively permanent structure, and in which sustained relations between the polarisation planes of the radiation and the expansion direction may be expected. Clearly, in any one object, the agreement could be entirely fortuitous, and the

Fig. 3 Radio linear polarisation properties of 3C 345: *a* the percentage polarisation at 3.7 cm, and *b*, the position angle (corrected for Faraday rotation) at 11.1 cm (○), 6 cm (×), and 3.7 cm (●).



crucial test will be to see if similar relationships can be found in other QSOs. It can also be said that the presence of long term systematic polarisation effects does not give support to QSO models that involve an array of unrelated sources.

I am grateful to Barbieri *et al.* and Cocke and Pacholczyk for sending me preprints before publication.

The Kitt Peak National Observatory is operated by the Association of Universities for Research in Astronomy, Inc., under contract with the NSF.

Received 14 March; accepted 22 April 1977.

¹ Dent, W. A. *Science* **148**, 1674 (1965).

² Goldsmith, D. H. & Kinman, T. D. *Astrophys. J.* **142**, 1693 (1965).

- ³ Kinman, T. D. *Astrophys. J. Lett.* **148**, L53 (1967).
⁴ Cohen, M. H. *et al. Astrophys. J. Lett.* **206**, L1 (1976).
⁵ Wittels, J. J. *et al. Astrophys. J. Lett.* **206**, L75 (1976).
⁶ Wittels, J. J. *et al. Astron. J.* **81**, 933 (1976).
⁷ Rees, M. J. *Nature* **211**, 46 (1966).
⁸ Dent, W. A. *Science* **175**, 1105 (1972).
⁹ Whitney, A. R. *et al. Science* **173**, 225 (1971).
¹⁰ Cavaliere, A., Morrison, P. & Sartori, L. *Science* **173**, 525 (1971).
¹¹ Sanders, R. H. *Nature* **248**, 390 (1974).
¹² Cocke, W. J. & Pacholczyk, A. G. *Nature* **265**, 608 (1977).
¹³ Elvius, A. *Lowell Obs. Bull.* No. 142 (Vol. VII, No. 5) (1968).
¹⁴ Visvanathan, N. *Astrophys. J.* **179**, 1 (1973).
¹⁵ Babadzhanjanz, M. K., Hagen-Thorn, V. A. & Semjonova, E. V. *Astrofizika* **11**, 385 (1975).
¹⁶ Kinman, T. D., Lamla, E., Ciurla, T., Harlan, E. & Wirtanen, C. A. *Astrophys. J.* **152**, 357 (1968).
¹⁷ Barbieri, C. & Abati-Erculiani, L. *Mem. Soc. astr. Ital.* **39**, 421 (1968).
¹⁸ Lu, P. K. *Astron. J.* **77**, 829 (1972).
¹⁹ Tritton, K. P. & Selmes, R. A. *Mon. Not. R. astr. Soc.* **153**, 453 (1971).
²⁰ Angione, R. J. *Astron. J.* **78**, 353 (1973).
²¹ Hayes, P. *Mon. Not. R. astr. Soc.* **173**, 553 (1975).
²² Altschuler, D. R. & Wardle, J. F. C. *Mem. R. astr. Soc.* **82**, 1 (1976).
²³ Seielstad, G. A. & Berge, G. L. *Astron. J.* **80**, 271 (1975).

A gating signal for the potassium channel?

R. H. Adrian & A. R. Peres

Physiological Laboratory, University of Cambridge

It is suggested that some part of the intramembrane charge movement in striated muscle is associated with the opening and closing of the potassium channel.

THE control of several membrane processes has been associated with small asymmetric displacement currents which may signal intramembrane charge displacement or dipole movement. Schneider and Chandler¹ reported a charge movement in striated muscle which they suggested could be associated with activation of contraction or possibly with activation of potassium conductance². Armstrong and Bezanilla³ and Keynes and Rojas⁴ have investigated the 'gating' currents associated with the activation of sodium conductance in the giant axon of the squid. We report here experiments on the charge movement in striated muscle (sartorius muscle, *Rana temporaria*) which show that it may be separated into two components; one has approximately first order kinetics, and the other has kinetics which seem to associate it with the potassium channel.

Measurement of charge movement

Sartorius muscles were examined in a solution designed to minimise all ionic currents and to prevent contraction. The solution was made hypertonic with sucrose and contained tetraethylammonium ion (TEA) and tetrodotoxin to eliminate sodium current and potassium current⁵. The composition of the solutions was Rb₂SO₄, 5 mM; TEA Cl, 15 mM; (TEA)₂SO₄, 80 mM; CaSO₄, 8 mM; sucrose, 350 mM; TTX, 10⁻⁷ M (ref.6). Three electrodes inserted close to the pelvic end of a fibre were used to apply exponentially blunted steps of potential ($\tau = 1$ ms, duration 128 ms) to the membrane at the end of the fibre¹. Records of the total current (I) injected into the fibre, the membrane current (i_m) as the potential difference (ΔV) between the two potential measuring electrodes ($i_m = 2\Delta V/3l^2r_i$, where l is the electrode separation), and the membrane potential (V) at a distance l from the end of the fibre were taken by digitalising the signals and filing them on the disk of a PDP 11 computer. Using this information appropriate programmes calculated the linear membrane constants and displayed nonlinear membrane currents on the computer oscilloscope. Photographic records were also taken of the analogue signals from the fibre to display the complete sequence of prepulses and measuring pulse. Membrane resistance and internal longitudinal resistance (r_i) were calculated from the steady-state values of I , V , and ΔV from a 10-mV control step. The membrane capacity was calculated from the internal

resistance (r_i) and the integral of the transient membrane current (ref. 6)

$$C_{\text{eff}} = \frac{2}{3l^2r_iV} \int_0^\infty \left[\Delta V_t(t) - \frac{\Delta V(\infty)}{V(\infty)} V(t) \right] dt \quad (1)$$

All experiments were done at 2–4 °C.

The detection of charge movement is best done by subtracting the membrane currents required for two identically shaped steps of potential. One of these steps, the test step, is in a range of membrane potential where charges or dipoles are moved by a change of potential: the other, the control step, is ideally in a range where the charges are not moved. For instance one can apply two pulses of different magnitude both starting from the same potential, a potential at which the charge movement is saturated. Alternatively one can apply two pulses of the same size but starting from different potentials. If the steps are not the same size the currents require appropriate scaling before subtraction. The subtraction removes currents in linear membrane elements; linear ionic currents and linear polarisation current in the membrane dielectric due to induced dipoles. Movement of charged molecules or dipole reorientation will in general be nonlinear if a sufficient potential range can be explored. Nevertheless, the subtraction method can only give the difference between the charge moved in test and control pulses, so that if charge moves in the control pulse because the pulses are at insufficiently separated potentials or because there are several species of charges moving the interpretation of nonlinear current may be difficult.

The central sets of records in Figs 1 and 2 are proportional to the non-linear membrane current ($i_m'(t)$) which is defined by

$$i_m'(t) = \frac{2}{3l^2r_i} \left[\Delta V_t(t) - \frac{V_t(\infty)}{V_c(\infty)} \Delta V_c(t) \right] \quad (2)$$

where the subscripts t and c refer to test and control potential steps. In Fig. 1 both control and test steps start at -90 mV: the control step is +10 mV and the positive going test step is variable. This procedure with four control steps and one test step is the same as that used by Adrian and Almers⁶ to study the charge movement they called Charge 1. In Fig. 2 the control pulse is +10 mV from -90 mV: the test pulse is +10 mV from a membrane potential displaced in a positive direction by

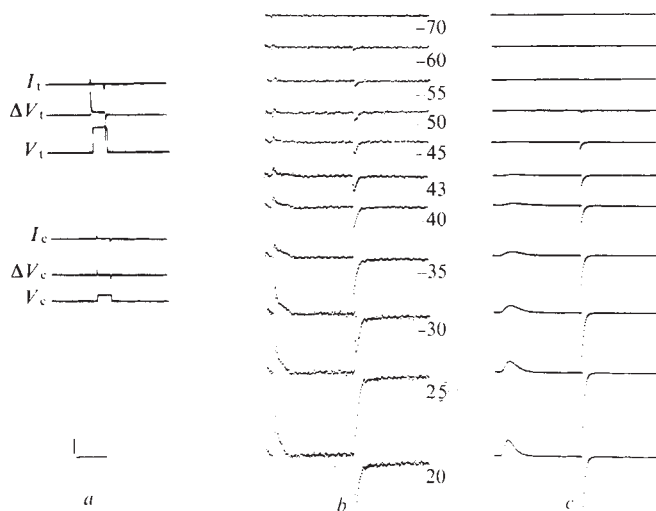


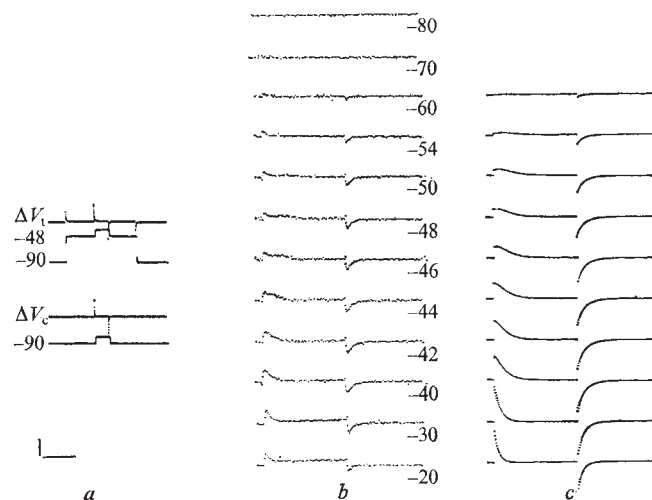
Fig. 1 *a*, Records of I , ΔV and V show a representative sequence of test and control pulses. *b* and *c* were generated by a PDP 11 computer and photographed from an oscilloscope. *b* Shows asymmetrical displacement currents

$$i'_m(t) \times \frac{3l^2 r_i}{2} = \left[(\Delta V_t(t) - \Delta V_c(t)) \frac{V_t(\infty)}{V_c(\infty)} \right]$$

produced by test pulses to an internal potential of -70 , -60 , -55 mV and so on. *c* Shows $d(n^4)/dt$ calculated for steps from -90 mV to -70 , -60 , -55 mV and so on (see text). Calibration lines for the computed traces are 50 ms for both traces; 10 mV for experimental traces and 0.033 ms^{-1} for $d(n^4)/dt$. Fibre 1501: l , 500 μm ; r_i , $21 \pm 0.7 \text{ M}\Omega \text{ cm}^{-1}$; R_m , $11.4 \pm 0.4 \text{ k}\Omega \text{ cm}^{-2}$; C_m , $6.0 \pm 0.1 \text{ }\mu\text{F cm}^{-2}$; diameter, $48 \pm 0.8 \text{ }\mu\text{m}$; temperature, 3.7°C . Fibre constants are means ($\pm$ s.e.m.) of five control steps of 10 mV at beginning of experiment.

a prepulse. In both Figs 1 and 2, the central column of records (*b*) shows the non-linear currents which result during and after the pulse when the potential range from -90 mV to 0 mV is explored. As has been reported previously^{2,8}, integration of the transient non-linear current for the 'on' and the 'off' of the step gives approximately equal charge movements, suggesting a capacitive origin for the current.

Fig. 2 As in Fig. 1, but the asymmetric displacement currents are generated by subtracting the ΔV records for two 10 mV steps starting from different potentials. Calibration lines are 50 ms; 3.3 mV for experimental records and 0.01 ms^{-1} for $d(n^4)/dt$. Fibre 1302: l , 500 μm ; r_i , $13.5 \pm 0.6 \text{ M}\Omega \text{ cm}^{-1}$; R_m , $8.7 \pm 0.3 \text{ k}\Omega \text{ m}^{-2}$; C_m , $5.6 \pm 0.2 \text{ }\mu\text{F cm}^{-2}$; diameter, $61 \pm 1.5 \text{ }\mu\text{m}$; temperature, 2.5°C .



In both Figs 1 and 2 the nonlinear current is most easily described in terms of two components. For increasing depolarisation during the test step (Fig. 1) or for increasing depolarisation preceding the test step (Fig. 2) one sees first a transient current which is maximal immediately following the 'on' of the step and decays more or less exponentially with a time constant of a few milliseconds. The reverse movement at the 'off' of the step is similarly exponential. At about -50 mV one begins to see a hump of current which occurs first in the middle of the step (that is maximum at about 60 ms from the beginning of the step). At the same potential there is a marked increase in the size of the exponentially decaying current at the 'off'. The hump of current is also clearly visible in the analogue record of ΔV (left hand column of Fig. 2). Over the next 10 mV of depolarisation the hump of current occurs earlier and becomes shorter and higher until eventually both 'on' and 'off' current transients decay more or less exponentially. Figure 1 shows the measured charge movements which take place between -90 mV and a series of more depolarised potentials: Fig. 2 shows the measured charge movements which take place in successive 10 mV intervals of the potential range -80 mV to -10 mV.

In both Figs 1 and 2 at membrane potentials between about -45 mV and -35 mV, the charge movement during the pulse can be described in terms of an exponentially decaying component together with a slower component which rises slowly and then falls. There is, however, no way to separate the components unambiguously and in particular no way of knowing whether, if there are two components, the initial value of the slowly developing movement is finite or zero. Nevertheless, it is evident that the slower charge movement during the voltage step is not well described by a first order process, though the movement after the voltage step might be.

The potassium conductance (g_K) is known to turn on with high order kinetics and off with first order kinetics^{7,9}. It seemed worthwhile, therefore, to see how far the charge movements in Figs 1 and 2 could be predicted by Hodgkin-Huxley kinetics¹⁰. In their description of the potassium conductance changes in the squid axon

$$g_K = \bar{g}_K n^4 \quad (3)$$

$$dn/dt = (1-n)\alpha_n - n\beta_n \quad (4)$$

$\bar{g}_K$ is a limiting conductance and α_n and β_n are rate constants depending only on the membrane potential (V). n^4 can be thought of as the probability that a potassium channel is in the open configuration. Adrian, Chandler and Hodgkin⁷ described the potassium conductance measured in sartorius muscle fibres at $1-3^\circ\text{C}$ with equations (3) and (4) and the rate constants at 3°C

$$\alpha_n = \frac{0.0044 (V+40)}{1 - \exp - [(V+40)/7]} \text{ ms}^{-1} \quad (5)$$

$$\beta_n = 0.0185 \exp - [(V+40)/40] \text{ ms}^{-1} \quad (6)$$

Figure 3 (which is reproduced from ref. 7) shows α_n , β_n , n_∞ (or n_0) and τ_n^{-1} where n_0 and n_∞ are given by

$$n_0 = \frac{\alpha_{n0}}{\alpha_{n0} + \beta_{n0}} \quad (7)$$

$$n_\infty = \frac{\alpha_{n\infty}}{\alpha_{n\infty} + \beta_{n\infty}} \quad (8)$$

and

$$\tau_n = \frac{1}{\alpha_{n\infty} + \beta_{n\infty}} \quad (9)$$

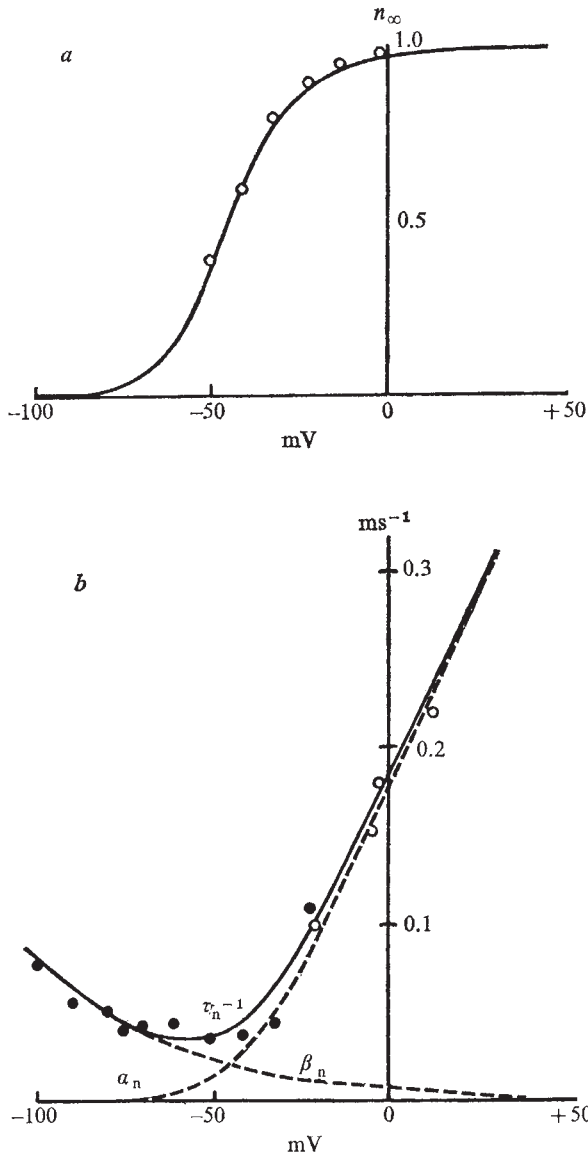


Fig. 3 *a*, Relation between n_{∞} (from steady-state potassium conductance) and membrane potential. Open circles represent experimental points; the line is calculated from

$$n_{\infty} = \frac{\alpha_n}{\alpha_n + \beta_n}$$

b, Relation between rate constants (τ_n^{-1}) obtained from records of potassium currents produced by depolarisation (○) and repolarisations (●) to different potentials: the line is calculated from

$$\tau_n^{-1} = \alpha_n + \beta_n$$

The dotted lines are calculated by equations (5) and (6). (Figure reproduced from ref. 7.)

For a step change of potential from V_0 to V_{∞} the solution to equation (4) is

$$n = n_{\infty} - (n_{\infty} - n_0) \exp(-t/\tau_n) \quad (10)$$

Hodgkin and Huxley¹⁰ suggested that movement of four independent particles onto a potassium channel precedes its opening. The movement of each n particle was described by equation (4). On this basis one would expect to detect an exponentially decaying charge movement (i_n) whenever the potential was changed abruptly which would be given by

$$i_n \propto (dn/dt) = \tau_n^{-1} (n_{\infty} - n_0) \exp(-t/\tau_n) \quad (11)$$

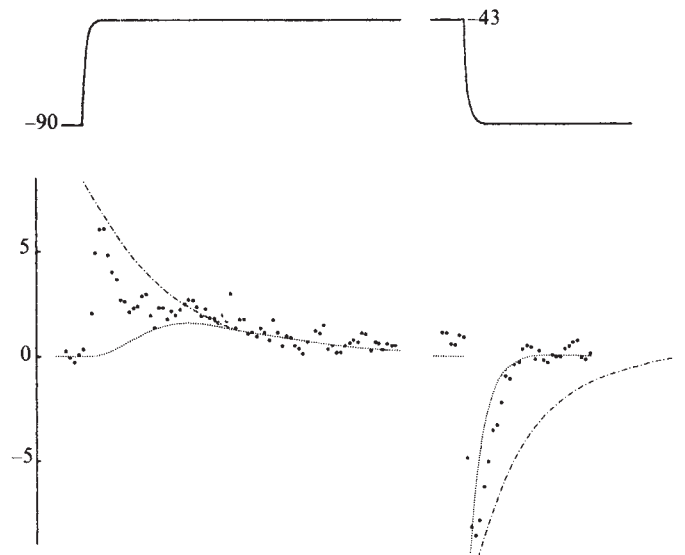
If the actual opening of a potassium channel involved the redistribution of other charges in the membrane field, or the entry of additional charges into the membrane, then in addition to the charge movement reflecting the preliminary and rate-limiting process there would be a charge movement associated with the opening process. This would resemble

$$\begin{aligned} \frac{d}{dt}(n^4) = & \frac{4(n_{\infty} - n_0)}{\tau_n} \left[n_{\infty}^3 \exp\left(\frac{-t}{\tau_n}\right) \right. \\ & - 3n_{\infty}^2 (n_{\infty} - n_0) \exp\left(\frac{-2t}{\tau_n}\right) + \\ & + 3n_{\infty} (n_{\infty} - n_0)^2 \exp\left(\frac{-3t}{\tau_n}\right) \\ & \left. - (n_{\infty} - n_0)^3 \exp\left(\frac{-4t}{\tau_n}\right) \right] \quad (12) \end{aligned}$$

Different kinetic models could be made to generate different combinations of exponential terms. Equation (12) by itself could not describe the total charge movement since it implies that the n process does not involve charge movement and therefore could not be potential dependent.

Figure 4 plots $i_m'(t)$ from Fig. 1 at a test potential of -43 mV and compares with it dn/dt and $d(n^4)/dt$ calculated by equations (5) to (12) for a potential step from -90 mV to -43 mV at a temperature of 4 °C (Q_{10} for rate constants was assumed to be 2.5). The two calculated curves have been arbitrarily scaled. It is clear that the charge during the test step could be described by an arbitrary combination of dn/dt and $d(n^4)/dt$. But it is equally clear that this would not be a unique description. It would, however, be difficult to describe the charge movement in terms of either function alone.

Fig. 4 The points are values of $[3I^2 r_i i_m'(t)]/2$ derived from a control step of +10 mV and a test step of +47 mV both from -90 mV. The dash-dot line is dn/dt , the dotted line $d(n^4)/dt$ both calculated from equations (5) to (12) at 4 °C. They are arbitrarily scaled so that the curves coincide in the latter part of their time course during the voltage step. Note that although the voltage step was blunted with a time constant of 1 ms, the calculated curves assume a sudden step. The points which are taken from the experiment in Fig. 1 are at 1-ms intervals on the horizontal axis; the vertical scale is in mV.



The right-hand columns of Figs 1 and 2 give $d(n^4)/dt$ calculated for the voltage steps of the experimental records. They show the extent to which later parts of the charge movements resemble the differential of n raised to a power greater than the first. In Fig. 1, $d(n^4)/dt = 0$ at the beginning of the step since n_0 at -90 mV is effectively zero. In Fig. 2 at the beginning of the 10 mV test step n_0 is not equal to zero and for $t = 0$

$$\frac{d}{dt}(n^4) = \frac{4n^3_0(n_\infty - n_0)}{\tau_n}$$

Figure 5 extends the similarity between part of the detected charge movement and $d(n^4)/dt$. In Fig. 5a and b charge is moved between the same potentials but in opposite directions: both charge movement and $d(n^4)/dt$ are approximately exponential when the potential change is negative, $d(n^4)/dt$ and part of the charge movement are humped when the potential change is positive. Figure 5c shows the charge movement when the potential changes from -32 to -42 mV: Fig. 5a shows the charge movement at the same potential but when the potential changes from -52 to -42 mV. Though the charge movement takes place at the same potential its time course depends on the direction of movement. The same kind of difference is seen in $d(n^4)/dt$.

Separation of charge movement into components

The question arises whether all or only part of the initial and approximately exponential charge movement seen in the experimental records should be associated with the more complex later charge movement. It would be difficult from the records to separate the charge movement into components. But in striated muscle it is known that the potassium conductance inactivates with time. This inactivation resembles, but is much slower than, the inactivation of the sodium conductance. Adrian, Chandler and Hodgkin¹⁷ have shown that the potassium conductance is about 70% inactivated when the membrane potential is held at -40 mV. If the membrane potential is held at -40 mV, but stepped to -90 mV for a 'control' +10 mV pulse and to other membrane potentials for 'test' 10 mV pulses, then the detected charge movement seems to consist of an early exponential component only. The voltage distribution of this moved charge (Q against V curve) does not resemble n_∞ being less steeply voltage dependent and with a transition potential at about -15 to -20 mV. The shape of this curve

resembles that postulated by Adrian, Chandler and Rakowski¹¹ to account for repriming of contraction after long depolarisation.

If the membrane potential is held at -20 mV both contraction and potassium conductance are inactivated and if charge movement is examined over the same potential range (-90 mV to 0 mV), both early exponential and late slow charge movements are absent. This effect of holding potential charge movement was first reported by Chandler, Rakowski and Schneider¹². It seems possible therefore that part, at least, of the early and exponential charge movement may be related to contraction and that the later slow charge movement as well as an undetermined part of the early more or less exponential charge movement may be related to the potassium channel.

In what follows we shall refer to this latter current as the opening signal or current, making the plausible but unproved assumption that it is a signal from the potassium channel. The implication is that external TEA blocks the potassium channel by combining with an external site on the channel^{5,13} but leaves unaltered the opening and closing of an internally located gate. The opening of this gate is associated with a movement of charge in or perhaps into the membrane: its closing with an equal and opposite movement of charge.

How many charges move when one channel opens? To estimate this quantity we need to know the conductance of a single potassium channel. We shall assume a value between 10^{-12} and 10^{-11} S (ref. 14). If the total opening charge is 10 nC μF^{-1} and the membrane capacity is 5 $\mu\text{F cm}^{-2}$ the charge is 50 nC cm^{-2} of fibre surface. When this amount of charge has moved, in the absence of TEA the potassium conductance would be 20 mS cm^{-2} (ref. 7). If the channel conductance is 4 pS (ref. 15) each opening event is associated with the movement of 60 charges. This figure is no more reliable than the rather uncertain quantities on which it is based, but it seems to exclude the possibility that the opening current is due to the entry of an internal TEA ion or other internal ion into the blocked channel when the gate opens¹⁶. Equally, though it might be consistent with models which involve the association of several multiply charged molecules in the opening process, the magnitude of the opening signal is unexpectedly large¹⁷.

In previous studies the charge movement in normally polarised muscle has been called charge 1 (refs 2, 8, 17). Chandler, Rakowski and Schneider^{2,12} gave careful consideration to an association between charge 1 and potassium activation but felt that there was insufficient evidence for a firm conclusion. Nonetheless, though comment has been made on the complexity of its kinetics¹⁸, charge 1 has been tentatively considered to be one species of charge and more likely to be associated with contraction activation than potassium channel activation¹⁷. Our results suggest that charge 1 may be two distinct charge species. Movement of one of these charges is an approximately first order process and this charge remains a candidate for being a regulator of contraction activation. The second charge could be associated with the potassium channel. The kinetics of this second charge movement favour an association with the opening of the potassium channel. But the size and complex time course argue against any simple link of the kind suggested by Hodgkin and Huxley¹⁰. Moreover, it is still possible that the detected charge movement is related to some high order process other than potassium gating.

We thank EMBO for the award of a Fellowship to enable A.P. to visit Cambridge, and the *Journal of Physiology* for permission to reproduce Fig. 3.

Received 23 February; accepted 2 May 1977.

¹ Schneider, M. F. & Chandler, W. K. *Nature* **242**, 224–246 (1973).

² Chandler, W. K., Rakowski, R. F. & Schneider, M. F. *J. Physiol., Lond.* **254**, 245–283 (1976).

³ Armstrong, C. M. & Bezanilla, F. *Nature* **242**, 459–461 (1973).

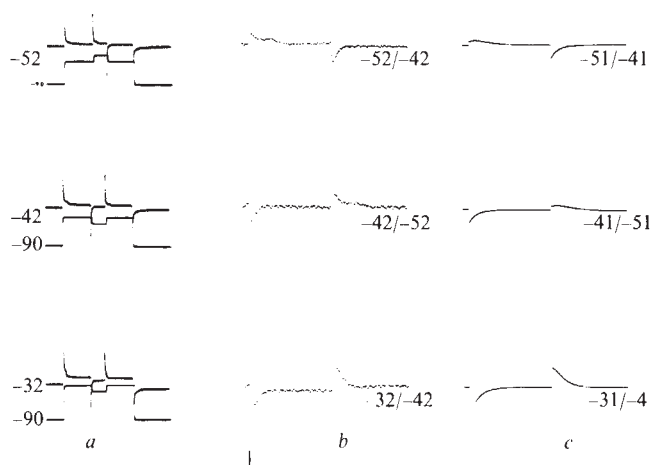
⁴ Keynes, R. D. & Rojas, E. *J. Physiol., Lond.* **233**, 28–30P (1973).

⁵ Stanfield, P. R. *J. Physiol., Lond.* **209**, 209–229 (1970).

⁶ Adrian, R. H. & Almers, W. *J. Physiol., Lond.* **254**, 317–338 (1976).

⁷ Adrian, R. H., Chandler, W. K. & Hodgkin, A. L. *J. Physiol., Lond.* **208**, 607–644 (1970).

Fig. 5 As in Fig. 2, but showing that configuration of displacement current depends on its direction. Calibration lines are 50 ms; 3.3 mV for experimental computer traces and 0.01 ms⁻¹ for $d(n^4)/dt$. Fibre 1303; l , 500 μm ; r_i , 11.8 ± 1.3 M Ω cm⁻¹; R_m , 14.3 ± 0.6 K Ω cm⁻²; C_m , 11.9 ± 0.2 $\mu\text{F cm}^{-2}$; diameter, 66 ± 3.2 μm ; temperature 2.6 °C.



- ⁸ Adrian, R. H. & Almers, W. *J. Physiol., Lond.* **254**, 339–360 (1976).
⁹ Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **116**, 473–496 (1952).
¹⁰ Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **117**, 500–544 (1952).
¹¹ Adrian, R. H., Chandler, W. K. & Rakowski, R. F. *J. Physiol., Lond.* **254**, 361–388 (1976).
¹² Chandler, W. K., Rakowski, R. F. & Schneider, M. F. *J. Physiol., Lond.* **254**, 285–316 (1976).
¹³ Armstrong, C. M. & Hillie, B. J. *gen. Physiol.* **59**, 388–400 (1972).
¹⁴ Armstrong, C. M. *Q. Rev. Biophys.* **7**, 179–209 (1974).
¹⁵ Begenisch, T. & Stevens, C. F. *Biophys. J.* **15**, 843–846 (1975).
¹⁶ Armstrong, C. M. *J. gen. Physiol.* **50**, 491–503 (1966).
¹⁷ Almers, W. *J. Physiol., Lond.* **262**, 613–637 (1976).
¹⁸ Almers, W., Adrian, R. H. & Levinson, S. R. *Ann. N.Y. Acad. Sci.* **264**, 278–292 (1975).

DNA complementary to parathyroid mRNA directs synthesis of pre-proparathyroid hormone in a linked transcription–translation system

Henry M. Kronenberg, Bryan E. Roberts*, Joel F. Habener, John T. Potts, Jr & Alexander Rich

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and Endocrine Unit, Massachusetts General Hospital, Boston, Massachusetts 02114

DNA complementary in sequence to the messenger RNA for pre-proparathyroid hormone was synthesised using reverse transcriptase. In a linked transcription–translation system using RNA polymerase and cell-free extract from wheat germ, the DNA directed the synthesis of a protein identified as pre-proparathyroid hormone by N-terminal sequencing and by electrophoretic and immunologic criteria.

PARATHYROID hormone (PTH) is a small polypeptide of 84 amino acids that is of critical importance in the regulation of calcium levels in the extracellular fluid of higher vertebrates (see ref. 1 for review). In the parathyroid gland, PTH is first synthesised as a precursor of 115 amino acids, pre-proparathyroid hormone (Pre-ProPTH) (ref. 2). During or shortly after synthesis *in vivo*, Pre-ProPTH undergoes two consecutive proteolytic cleavages, resulting in the formation of PTH through an intermediate precursor, proparathyroid hormone, a polypeptide of 90 amino acids³. Both proteolytic cleavages occur at the amino-terminal end of the precursors. The primary structures of PTH and the two biosynthetic precursors have been determined^{4–6}.

Previously we have prepared and characterised the messenger RNA for Pre-ProPTH (refs 4, 5, 7). Using reverse transcriptase, we have now synthesised DNA complementary in sequence (cDNA) to this message. Recently, several investigators^{8–10} have developed conditions for using reverse transcriptase to synthesise cDNA molecules as large as their mRNA templates. We have used similar conditions and report here that the cDNA is large enough to direct the synthesis of Pre-ProPTH in a linked transcription–translation cell-free system. The cDNA was transcribed using wheat germ RNA polymerase, and the newly synthesised RNA was then used to direct protein synthesis in a cell-free extract. Control experiments showed that this synthesis was not due to contamination of the cDNA by the initial mRNA. The Pre-ProPTH synthesised was characterised by molecular sizing on polyacrylamide gels, specific reactivity with antiserum to PTH, and amino-acid sequencing at the N terminus.

DNA-directed Pre-ProPTH synthesis

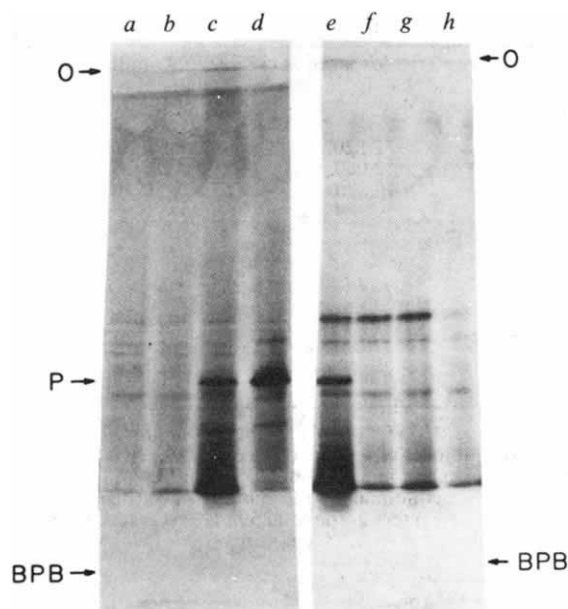
Complementary DNA was synthesised from partially purified Pre-ProPTH mRNA using conditions similar to

those described by Efstratiadis *et al.*⁸ (see legend to Fig. 1). Electrophoresis in a 5% polyacrylamide–98% formamide gel showed that the cDNA is heterogeneous in size, with a major species containing almost 800 nucleotides. Its characterisation will be described more fully elsewhere. The Pre-ProPTH mRNA was not characterised in detail here, but earlier studies indicated that it is approximately the same size as the cDNA (refs 4, 6).

We tested the ability of the cDNA prepared from Pre-ProPTH mRNA to direct the synthesis of Pre-ProPTH in the linked transcription–translation system. After treatment with alkali to hydrolyse the RNA, the cDNA stimulated the synthesis of a protein that co-migrated with Pre-ProPTH on a SDS–urea, 15–20% polyacrylamide gradient gel (Fig. 1c, d). The synthesis of this protein was dependent on the addition of wheat germ RNA polymerase (Fig. 1b). Moreover, all of the proteins synthesised in the absence of RNA polymerase co-migrated with those that were synthesised by the wheat germ extract in the absence of any added nucleic acid (Fig. 1a, b).

The incorporation of ³H-UTP into trichloroacetic-acid-precipitable material was dependent upon the addition of RNA polymerase to the cDNA (Table 1). It was reasonable to conclude that this newly-synthesised RNA was translated into Pre-ProPTH, but other possible alternatives had to be examined. For example, the polymerase might have stimulated the synthesis of non-functional RNA that protected from degradation some of the original mRNA present as a contaminant in the cDNA preparation. Therefore, experiment 2 in Fig. 1 was performed in the presence of carrier transfer RNA (1 mg ml⁻¹ during the transcription), and the results make this explanation unlikely. The synthesis of Pre-ProPTH required the presence of RNA polymerase (Fig. 1e, g), even though carrier tRNA was included in each incubation. Alternatively, the polymerase might have removed contaminating mRNA from a cDNA–mRNA hybrid, allowing the mRNA to bind ribosomes and undergo translation. These possibilities and most other explanations that postulate contamination of the cDNA preparation with mRNA were eliminated by treating the cDNA preparation with DNase. Complementary DNA treated with pancreatic DNase during the transcription did not stimulate protein synthesis in the linked transcription–translation system (Fig. 1e, f). The possibility that the effect of DNase was not due to digestion of cDNA, but rather due to the action of contaminating RNase was excluded. In the conditions of the incubations, mRNA added separately to the reaction with DNase retained 84% of its original activity (Table 1).

*Present address: Department of Biology, Brandeis University, Waltham, Massachusetts.



contained $2 A_{260}$ units ml^{-1} Pre-ProPTH mRNA, four deoxyribonucleotide triphosphates, 1 mM each, and ^3H -dCPT adjusted to 0.4 Ci mmol^{-1} . Avian myeloblastosis virus reverse transcriptase was kindly supplied by J. W. Beard. After 3 h at 43°C , the reaction mixture was adjusted to 0.1 N in NaOH and incubated for another 16 h at 37°C . After neutralisation of the NaOH with an equal amount of HCl and adjustment to 0.1 M in Tris-HCl ($\text{pH } 9.0$), the nucleic acid was extracted once with an equal volume of phenol-chloroform-isoamyl alcohol and passed over a 5 ml Sephadex G1-50 column in H_2O . Fractions eluting in the void volume of the column were pooled, adjusted to 0.1 M NaCl and precipitated with 2 volumes of ethanol. In experiment 2, but not in experiment 1, a mixture of purified wheat germ tRNA provided by T. Fraser, was added at $50 \mu\text{g ml}^{-1}$ as a carrier before the ethanol precipitation. The conditions used for the linked transcription-translation reactions were similar to those described previously^{11, 12} except that the transcription reaction was carried out for 1 h instead of 15 min. Each reaction mixture contained approximately $0.1 \mu\text{g}$ DNA, estimated from the specific activity of ^3H -CTP. In experiment 2 (channels (e)-(h)), assuming 100% recovery, the tRNA added as carrier was present at a concentration of 1 mg ml^{-1} during transcription and 0.2 mg ml^{-1} during translation. Addition of larger amounts of wheat germ tRNA to the reaction mixture inhibited translation. Different preparations of cDNA were used in each of the two experiments. ^{35}S -methionine was used as the radioactive amino acid. After the translation reaction was completed, the proteins in $5 \mu\text{l}$ of each $50 \mu\text{l}$ extract were analysed by electrophoresis on 15–20% polyacrylamide, SDS-urea gels²⁴ followed by fluorography²⁵. Gels were exposed to Kodak SB-5 film for 3 d at -70°C . Experiment 1: channel a, no DNA; b, no wheat germ RNA polymerase added; c, cDNA and RNA polymerase added; d, $0.04 \mu\text{g}$ mRNA added directly to $50 \mu\text{l}$ of the translation reaction mixture. Experiment 2: channel e, cDNA, RNA polymerase and tRNA added; f, same as (e) plus pancreatic DNase 0.1 mg ml^{-1} added during transcription; g, same as (e) without RNA polymerase added; h, same as (e) without cDNA or tRNA. O, origin; P, PreProPTH; BPB, bromophenol blue dye front.

Characterisation of protein synthesised in the linked system

We took advantage of the detailed characterisation of Pre-ProPTH (refs 4–6) to investigate the structure of the protein synthesised from cDNA. Immunoprecipitation of the products from the linked transcription-translation reaction with guinea pig antiserum to PTH and rabbit anti-guinea pig serum produced a polypeptide that co-migrated on sodium dodecyl sulphate (SDS)-urea polyacrylamide gels with authentic, mRNA-directed Pre-ProPTH (Fig. 2). Moreover, no ^{35}S -methionine-labelled peptide comigrated with Pre-ProPTH in immunoprecipitates prepared from cell-free reactions lacking either cDNA or RNA polymerase. The cDNA-directed product was not precipitated when treated with non-immune serum and rabbit anti-guinea pig serum or when bovine PTH ($35 \mu\text{g ml}^{-1}$) was included before the addition of anti-PTH antiserum and rabbit anti-guinea pig serum. The specificity of the immunological studies was strengthened further by control experiments in which anti-PTH antiserum and rabbit anti-guinea pig serum did not precipitate any of the numerous ^{35}S -methionine polypeptides synthesised in the linked system in response to SV40 DNA, form I (ref. 12). Thus, our cDNA-directed product contains antigenic determinants that are present in PTH.

Methionine is present at positions 1, 2, 7, 11 and 14 in the amino-terminal region of Pre-ProPTH (ref. 4). We therefore used Edman degradation to determine the location

of the methionines in the structure of the cDNA-directed product that was labelled with ^{35}S -methionine (Fig. 3). ^{35}S -methionine was released during the degradations at cycles 1, 2, 7, 11 and 14, a pattern identical to that found in a similar sequence analysis of mRNA-directed Pre-ProPTH. Thus, sequence studies, immunological data, and electrophoretic analysis indicate that the structure of the cDNA-directed product is very similar, if not identical, to the structure of Pre-ProPTH. We conclude that the cDNA that we have synthesised contains a faithful representation of the structural information for Pre-ProPTH and can express this information in the linked, transcription-translation system.

Discussion

The linked transcription-translation system provides an important method for assessing the fidelity and completeness of the cDNA products of the reverse transcription of mRNA. Furthermore, the method is independent of hybridisation and sizing tests. Figure 4 summarises the reactions required to synthesise Pre-ProPTH in the linked transcription-translation system. The synthesis of cDNA is presumed to begin at or near the 3'-terminus of the mRNA with an oligo-dT primer and clearly extends to the nucleotides encoding the N-terminal methionine of Pre-ProPTH. Therefore, the cDNA must include the complement of any untranslated region at the 3' end of the mRNA (region C in Fig. 4). It must also span the region coding

Table 1 cDNA-directed RNA and protein synthesis in the linked transcription-translation system

Experimental conditions	Trichloroacetic-acid insoluble radioactivity	
	RNA- ³ H c.p.m.	Protein- ³⁵ S c.p.m. × 10 ⁻³
Experiment 1		
<i>a</i> No nucleic acid	2,320	226
<i>b</i> No RNA polymerase	2,060	266
<i>c</i> cDNA-directed product	10,820	1,010
<i>d</i> mRNA-directed product		538
Experiment 2		
<i>e</i> cDNA-directed product		599
<i>f</i> DNase (0.1 mg ml ⁻¹) included along with cDNA		240
<i>g</i> No RNA polymerase		240
<i>h</i> No nucleic acid added		202
<i>i</i> mRNA (0.2 µg) added at time of transcription		1,447
<i>j</i> As (<i>i</i>) plus DNase (0.1 mg ml ⁻¹)		1,231

Experimental conditions detailed in Fig. 1 legend. RNA synthesis was measured using ³H-UTP (0.1 Ci mmol⁻¹). 0.5 µl of the 10 µl transcription reaction was counted before the addition of translation reagents. The values for radioactivity refer to the total (10 µl) in the reaction mixture. The RNA radioactivity in reactions (*b*) and (*c*) was derived by subtracting the known amount of ³H-cDNA (8,200 c.p.m.) added to each reaction from the total radioactivity in the reactions (10,260 and 19,020 respectively). ³⁵S-methionine (460 Ci mmol⁻¹) was used in the translation reactions. 1 µl of the 50 µl translation reaction was measured. The radioactivity in the Table refers to the total 50 µl. The reactions (*a*)–(*h*) correspond to reactions (*a*)–(*h*) in Fig. 1.

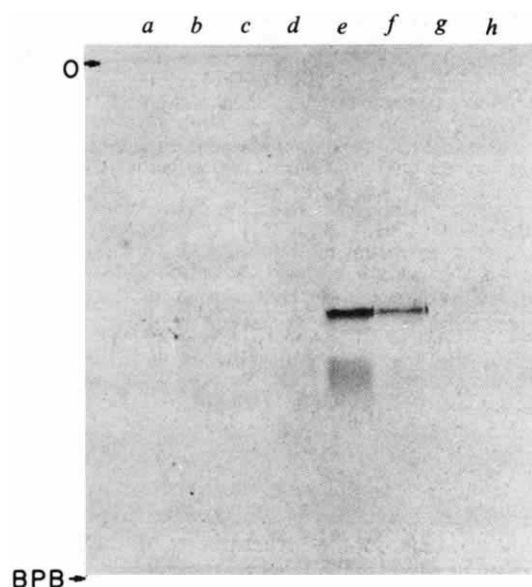


Fig. 2 Indirect immunoprecipitation of Pre-ProPTH. Approximately 100,000 c.p.m. of ³⁵S-methionine-labelled protein from each reaction were added to 200 µl 0.05 M veronal buffer (pH 8.5), 1% foetal bovine serum. Guinea pig anti-PTH serum (GPI) (ref. 26), diluted to a final concentration of 1:1000, was added to each tube except (*d*). After 24 h at 4 °C, 50 µl of rabbit anti-serum to guinea pig globulin (Arnel) were added to each tube, and the tubes were incubated at 4 °C for an additional 24 h. The resulting precipitates were collected by centrifugation at 1200g for 30 min, washed twice with 2 ml of 0.15 M NaCl, 0.02 M Tris-HCl (pH 7.5) and dried. The proteins were analysed by electrophoresis on an SDS-15–20% polyacrylamide slab gel²⁷. The gel was fluorographed²⁸ (channel *f* for 1 d, all the rest for 10 d) at -70 °C using Kodak SB-5 film. Immunoprecipitates of the following reactions are displayed: channel *a*, no cDNA or RNA polymerase added; *b*, cDNA added without RNA polymerase; *c*, cDNA and RNA polymerase added and non-radioactive PTH (35 µg ml⁻¹ during immune precipitation) added along with anti-PTH serum; *d*, cDNA and RNA polymerase added and non-immune serum substituted for anti-PTH serum at the same dilution; *e*, cDNA and RNA polymerase added; *f*, 0.08 µg mRNA added directly to wheat-germ translation reaction; *g*, 0.1 µg SV40 DNA (form I), and RNA polymerase added; *h*, same as *g* but non-radioactive PTH (35 µg ml⁻¹ during immune precipitation) added along with anti-PTH antiserum. O, origin; BPB, bromophenol blue front.

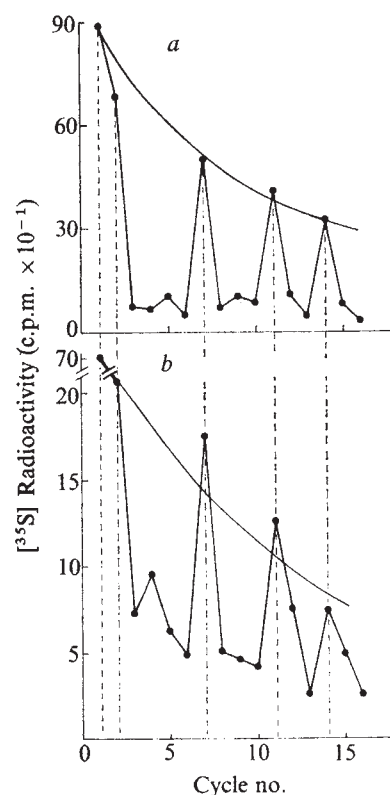


Fig. 3 Sequence analyses of *a*, ³⁵S-methionine-labelled Pre-ProPTH synthesised in response to parathyroid mRNA in the wheat-germ cell-free system and *b*, the major ³⁵S-methionine-labelled protein synthesised in the linked transcription-translation system in response to parathyroid cDNA. The Pre-ProPTH (*a*) was isolated by electrophoresis on an SDS-urea polyacrylamide gel²⁴ and the cDNA-directed protein by indirect immunoprecipitation with an antiserum to PTH. The radiolabelled proteins were subjected to automated Edman degradation (Beckman 890 sequencer) in the presence of 2 mg of apomyoglobin added as carrier. Aliquots of one-third of the total radioactivity released at each cycle of degradation were assayed directly for radioactivity as the butyl chloride extract. The identity of the radioactive amino acids was confirmed by chromatography of the remainder of each fraction on thin layer plates as described previously³. The continuous curved line represents the theoretical recoveries of methionine on each cycle based on the efficiency of degradation³.

for Pre-ProPTH (region B), and reach at least the initiator codon at the 5' end of the mRNA.

If region A contains sequences necessary for message

activity in the linked transcription-translation system, then our cDNA must extend into region A. Unfortunately, there is no firm information about the absolute requirement for

this region in our experiments. Shine and Dalgarno¹³ have shown that prokaryotic mRNA's contain sequences that are complementary to a sequence on 16S ribosomal RNA. Binding studies¹⁴ using mRNA from R17 bacteriophage and *Escherichia coli* ribosomes suggest that this sequence is functionally important. Ptashne *et al.*¹⁵ have suggested, however, that ribosomal-binding sequences on mRNA may not be absolutely required for mRNA activity. They have demonstrated that the bacteriophage λ repressor mRNA transcribed from the promoter p_{RM} does not require a 5'-terminal extension beyond the initiator AUG-codon for translation, albeit inefficient translation, *in vivo*. Furthermore, the functional importance of ribosomal-binding sequences¹⁶ or other sequences 5'-proximal to the initiator nucleotides has not been determined for eukaryotic messages. Thus we cannot be certain about the completeness of the polynucleotide sequence of our cDNA in region A.

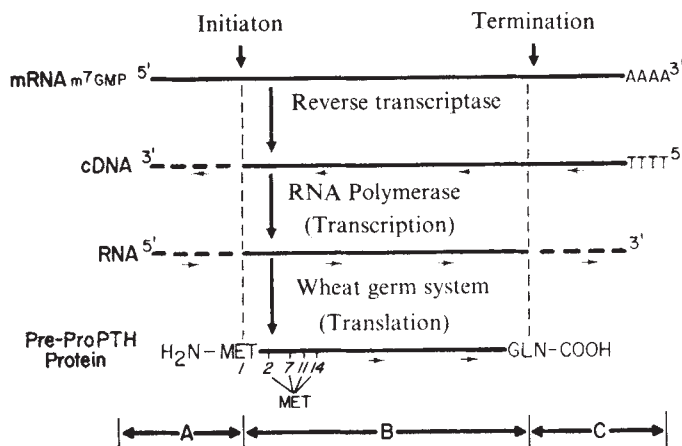


Fig. 4 Diagram of cDNA synthesis and synthesis of Pre-ProPTH in the linked *in vitro* transcription-translation system. Horizontal solid lines indicate sequences necessarily synthesised in the functioning linked system. Dotted lines indicate sequences that may have been synthesised. The N-terminal methionines of the Pre-ProPTH synthesised in the linked system are indicated. Horizontal arrows under cDNA, RNA, and Protein denote the direction of synthesis of these molecules.

Our demonstration that RNA synthesised in the linked transcription-translation system is accurately translated into Pre-ProPTH raises interesting questions regarding the role of the 7-methyl guanosine terminal in the activity of parathyroid mRNA. Kemper¹⁷ has reported that treatment of Pre-ProPTH mRNA sequentially with periodate and aniline inactivates the message almost completely. Control studies suggested that periodate and aniline were effective at least partly because they removed 5'-terminal 7-methyl guanosine¹⁷. These studies of Kemper and more recent data from Rose and Lodish¹⁸ demonstrate that periodate and aniline treatment can inactivate messages by producing alterations in the mRNA other than removal of 7-methyl guanosine. Therefore, whether or not the Pre-ProPTH mRNA absolutely requires a 7-methyl guanosine terminus in order to be translated is unresolved. In our studies either 'uncapped' RNA is translated or a 7-methyl guanosine cap is added to the RNA by the linked transcription-translation system. Analysis of the newly-synthesised RNA should resolve this issue.

We are as yet uncertain as to whether or not our functional cDNA is single or double-stranded. Using conditions similar to those that we have followed, Efstratiadis *et al.*⁸ found that only 9.2% of their globin cDNA was double stranded, as defined by resistance to the S1 nuclease, which is specific for single-stranded nucleic acid. This fraction became even smaller (4.4%) when actinomycin D

was added during reverse transcription. We did not include actinomycin in our reaction mixture during synthesis of cDNA and cannot be sure how much of our DNA is double-stranded. Conceivably, even a small amount of double-stranded DNA could be responsible for all meaningful transcription. Further studies using actinomycin D and techniques for separating single and double-stranded DNA are underway.

The methodology described here should be particularly useful for investigators attempting to introduce functional DNA into bacteria²⁰⁻²². At each stage of synthesis and modification of the DNA, the DNA can be monitored to assure that the nucleic acid retains sufficient information to code for appropriate proteins. After the vector has been successfully introduced into bacteria, restriction fragments of the recombinant DNA can be similarly assayed. If the particular protein of interest is not synthesised by the bacteria, the linked system should allow the detection of potentially functional sequences not properly transcribed by the bacteria. An additional strength of this methodology is that the cDNA employed in the studies need not be pure. Crude cDNA mixtures, which can later be purified by cloning, can be shown to contain the structural information for the proteins of interest.

The parathyroid cDNA is a suitable starting material for nucleotide sequence analysis²⁸. These data should be especially interesting in regions A and C of Fig. 4. Further, the synthesis of double-stranded parathyroid DNA should provide a substrate for detailed restriction endonuclease analysis. The ligation of specific DNA restriction fragments could generate DNA molecules with characteristic deletions and insertions. The behaviour of these modified DNA's in the linked transcription-translation system should allow a detailed analysis of the function of specific untranslated nucleotide sequences. Similar experiments can be carried out in the coding region B to create modified versions of PTH whose activities could be assayed in biological systems.

We thank R. Storti, D. Coen and S. Tomlinson for advice and assistance and M. Ernst for technical help. This work was supported by the US NIH, the John A. Hartford Foundation, the American Cancer Society. H.M.K. is a recipient of a US PHS Research Service Award and J.F.H. is an investigator of the Howard Hughes Medical Institute.

Received 7 March; accepted 19 April 1977.

- Potts, J. T. & Deftos, L. J. in *Duncan's Diseases of Metabolism* 2 (eds Bondy, P. K. & Rosenberg, L. E.) 1225-1430 (W. B. Saunders, Philadelphia, 1974).
- Habener, J. F., Potts, J. T., Jr & Rich, A. *J. biol. Chem.* 251, 3893-3899 (1976).
- Jacobs, J. W., Kemper, B., Niall, H. D., Habener, J. F. & Potts, J. T., Jr *Nature* 249, 155-157 (1974).
- Kemper, B., Habener, J. F., Ernst, M. D., Potts, J. T., Jr & Rich, A. *Biochemistry* 15, 15-19 (1976).
- Kemper, B., Habener, J. F., Potts, J. T., Jr & Rich, A. *Biochemistry* 15, 20-25 (1976).
- Habener, J. F., Kemper, B., Rich, A. & Potts, J. T., Jr *Recent Prog. Hormone Res.* 33, (in the press).
- Kemper, B., Habener, J. F., Mulligan, R. C., Potts, J. T., Jr & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* 71, 3731-3735 (1974).
- Efstratiadis, A., Maniatis, T., Kafatos, F., Jeffrey, A. & Vournakis, J. *Cell* 4, 367-378 (1975).
- Kacian, D. L. & Myers, J. C. *Proc. natn. Acad. Sci. U.S.A.* 73, 2191-2195 (1976).
- Monahan, J. J., Harris, S. E., Woo, S. L., Robberson, D. L. & O'Malley, B. W. *Biochemistry* 15, 223-233 (1976).
- Rosenblatt, S., Mulligan, R. C., Gorecki, M., Roberts, B. E. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* 73, 2747-2751 (1976).
- Roberts, B. E., Gorecki, M., Mulligan, R. C., Danna, K. J., Rosenblatt, S. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* 72, 1922-1926 (1975).
- Shine, J. & Dalgarno, L. *Nature* 254, 34-38 (1975).
- Steitz, J. A. & Jakes, K. *Proc. natn. Acad. Sci. U.S.A.* 72, 4734-4738 (1975).
- Ptashne, M. *et al. Science* 194, 156-161 (1976).
- Dasgupta, R., Shih, D. S., Saris, C. & Kaesberg, P. *Nature* 256, 624-628 (1975).
- Kemper, B. *Nature* 262, 321-323 (1976).
- Rose, J. K. & Lodish, H. F. *Nature* 262, 32-37 (1976).
- Maniatis, T., Kee, S. G., Efstratiadis, A. & Kafatos, F. *Cell* 8, 163-182 (1976).
- Higuchi, R., Paddock, G. V., Wall, R. & Salser, W. *Proc. natn. Acad. Sci. U.S.A.* 73, 3146-3150 (1976).
- Rabbitts, T. H. *Nature* 260, 221-225 (1976).
- Rougeon, F., Kourilsky, P. & Mach, B. *Nucleic Acids Res.* 2, 2365-2378 (1975).
- Anderson, C. W., Lewis, J. B., Atkins, J. F. & Gesteland, R. F. *Proc. natn. Acad. Sci. U.S.A.* 71, 2756-2760 (1974).
- Storti, R. V., Coen, D. M. & Rich, A. *Cell* 8, 251-258 (1976).
- Laskey, R. A. & Mills, A. D. *Eur. J. Biochem.* 56, 335-341 (1975).
- Segre, G. V., Habener, J. F., Powell, D., Tregear, G. W. & Potts, J. T., Jr *J. clin. Invest.* 51, 3163-3172 (1972).
- Laemmli, U. K. *Nature* 227, 680-685 (1970).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* 74, 560-564 (1977).

Mechanism of acid protease catalysis based on the crystal structure of penicillopepsin

Michael N. G. James, I-Nan Hsu & Louis T. J. Delbaere

MRC Group on Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta, T6G 2H7, Canada

A proposed mechanism for the catalytic hydrolysis of peptide bonds by acid proteases is similar in many respects to the Zn-carbonyl mechanism previously derived for carboxypeptidase A. In the acid proteases the electrophilic component is the proton shared by Asp-32 and Asp-215; Tyr-75 donates its proton to the amide nitrogen of the scissile bond and an OH⁻ ion from a water molecule bound between the carboxyl group of Asp-32 and the substrate attacks the carbonyl carbon atom.

ACID proteases catalyse the hydrolysis of peptide bonds at acidic pH values. The solutions of high resolution structures of three of these enzymes isolated from diverse microbial sources have been reported^{1,2}. These crystal structures will be of fundamental importance in elucidating the catalytic mechanism of acid proteases. The electron density map of penicillopepsin¹ is exceptionally well defined and the region of the active site is unambiguous in its interpretation. In addition, this is the only acid protease for which both the primary structure and crystal structure have been determined. We are therefore in a position to propose a mechanism based on structural information and previously documented chemical data.

Our primary concern in the present discussion is the bond-breaking step of polypeptide substrate hydrolysis by acid proteases. Further studies with substrate analogues, inhibitors and virtual substrates will add to our present interpretation of the catalytic mechanism and should provide the details of the enzyme-substrate interactions of elongated substrate peptide chains.

The history of studies of the catalytic mechanism has been most recently reviewed by Fruton³. In that article a number of proposals for future research were also made. We present here additional data which could help in the interpretation of those experiments. Moreover, neither a covalent amino-enzyme nor a covalent acyl-enzyme intermediate is required by the mechanism described here. The existence of these covalent intermediates was essential to many of the previous hypotheses but the intermediates have not been detected.

Crystal data

Penicillopepsin crystallises in the monoclinic system (space group, C2, with unit cell dimensions $a = 97.38(8)$, $b = 46.62(3)$, $c = 65.39(6)$ Å and $\beta = 115.4(3)^\circ$) from a 9 mg ml⁻¹ solution of the enzyme and 2.5 M Li₂SO₄, 0.1 M sodium acetate at pH 4.4. The low-resolution crystal structure at 6 Å showed that the molecule was decidedly asymmetric with dimensions

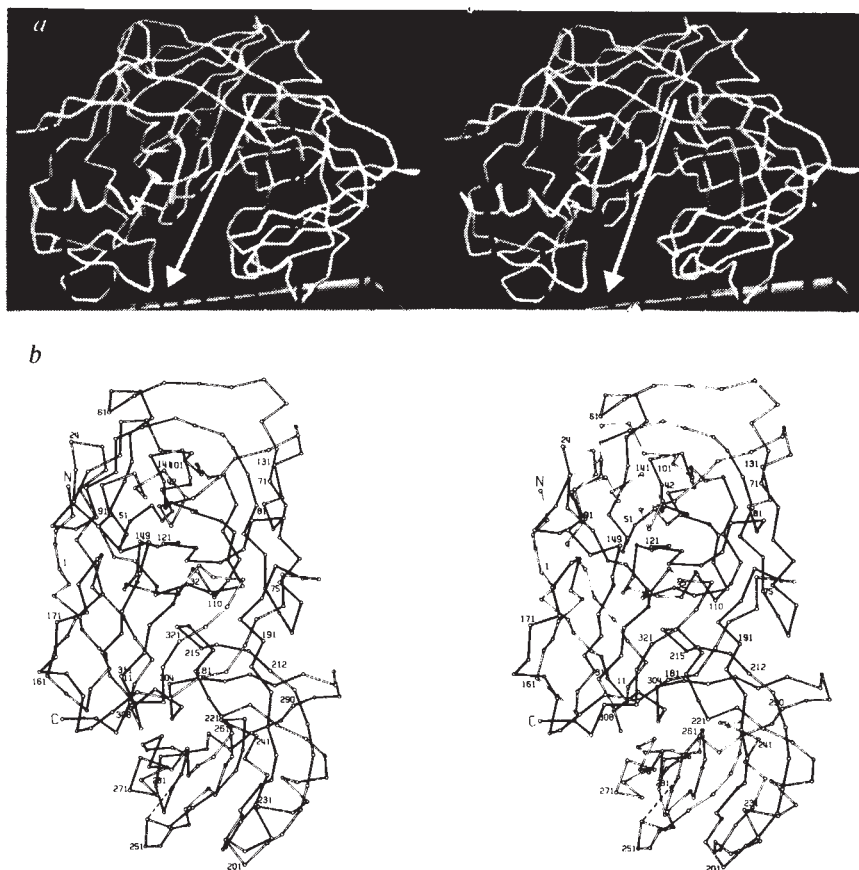


Fig. 1 *a*, Stereoscopic representation of the penicillopepsin α -carbon model. The arrow is shown in the extended substrate binding cleft which divides the molecule into two approximately equal halves (the N-terminal domain is to the right, the C-terminal domain to the left). Asp-32 (in the N-terminal domain) and Asp-215 (in the C-terminal domain) have their α -carbon positions in a dark colour. *b*, Stereoscopic drawing (from the program ORTEP of Dr Carroll Johnson) of the polypeptide chain. The view is from the top of that shown in (*a*). Every tenth α -carbon atom has been assigned a number corresponding to the sequence given in ref. 1. The active site region is that region shown with black connectivity; Asp-32 and Asp-215 are labelled.

$65 \times 49 \times 39 \text{ \AA}$ (ref. 4). Intensity data to $2.8 \text{ \AA}$ resolution (6,700 unique reflections) were obtained from a single crystal of native penicillopepsin and each of the eight heavy-atom derivatives used for determining phases by the method of multiple isomorphous replacement (MIR). The data collection protocol is discussed in more detail in refs 5 and 6 and the statistics regarding the phase determination are given in ref. 1. The overall mean figure of merit of the 6,127 statistically significant reflections was 0.90 (for zero error in phase angle this index has a value of 1.0). Observed native structure factor amplitudes and the 'best' phases⁷ were used to compute the electron density function of the penicillopepsin molecule. This map was interpreted using the partial sequence data⁸ for 16 fragments of penicillopepsin. In this way, it was possible to correlate the majority of the 322 residues in the polypeptide chain with the sequence data. Eight amino acid residues, however, were assigned from electron density distribution alone¹.

Structure and function

Figure 1 shows two different representations of the α -carbon model of penicillopepsin. A deep cleft separates the two lobes of the molecule and the active site is located within this cleft. A more detailed description of the polypeptide chain conformation is given in ref. 1. The positions of Asp-32 and Asp-215 are marked in the figures as an aid to orientation. The extensive hydrophobic core (in the shape of a croissant) appears as a relatively 'empty' region surrounded by the 18-strand mixed β -sheet, the major structural determinant of penicillopepsin. This hydrophobic core is highly ordered in the electron density map and the aromatic and aliphatic side chains are absolutely unambiguous in their assignment.

1,2-Epoxy-3-(*p*-nitrophenoxy) propane binding to penicillopepsin

The reagent 1,2-epoxy-3-(*p*-nitrophenoxy) propane (EPNP) is a specific inhibitor of acid proteases and has been used to identify the carboxylate group of Asp-32 at the active site⁹. Direct proof that Asp-32 and Asp-215 are present in the active site of penicillopepsin comes from the following experiment. A single crystal of penicillopepsin was soaked for 15 d in a saturated solution of EPNP in standard buffer at pH 4.4. The $2.8\text{-}\text{\AA}$ diffraction intensity data were collected from this crystal and a difference electron density map was computed from the amplitudes ($|F_{\text{EPNP}}| - |F_p|$) and the MIR phases (α_p). This map had significant density in only two regions. One region corresponds to what we interpret as a non-covalently-bound EPNP molecule in a solvent cavity between two protein molecules in the vicinity of residues 27–29 and 116–117. The second and more interesting region involves residues in the active site. This portion of the difference map is shown in Fig. 2. The positive values of difference density, indicated by solid lines, show that two molecules of EPNP have reacted covalently with penicillopepsin, one with the carboxyl group of Asp-32, the other with the carboxyl group of Asp-215. In addition, Fig. 2

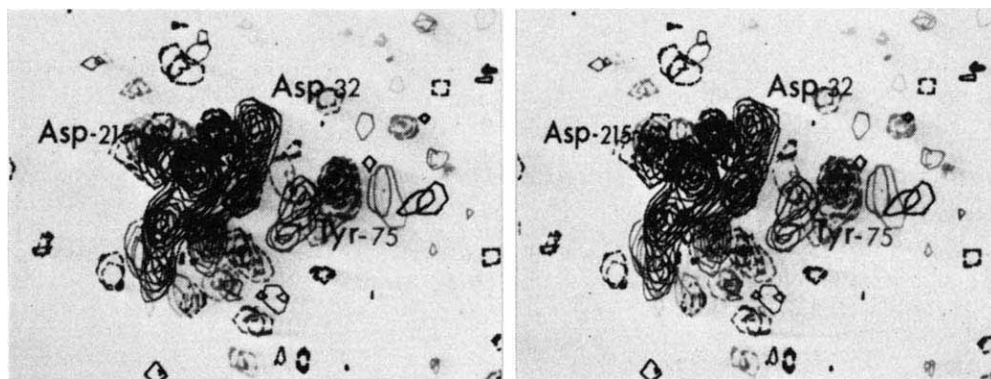
shows a concomitant pronounced conformational change of Tyr-75 in which the side chain phenoxyl group moves from a position where it was hydrogen-bonded to Trp-39 (an invariant residue in penicillopepsin and pepsin) to one in which it lies coplanar with the *p*-nitrophenoxy moiety of the EPNP molecule bound to Asp-32.

Room temperature studies of Chen and Tang¹⁰ on the inhibition of porcine pepsin with EPNP resulted in two peptides which had covalently bound EPNP; one with the sequence Asp(EPNP)–Thr–Gly–Ser–Ser–Asn (residues 32–37) and the other with sequence Ile–Val–Asp(EPNP) (residues 213–215). A similar experiment at 4°C by Hartsuck and Tang⁹ resulted in only one EPNP reactive peptide which corresponded to the Asp-32 sequence. Chen and Tang¹⁰, perhaps in view of the low temperature results, chose to assign the EPNP reactive site only to Asp-32. Their observations and our present knowledge of the sequence are in good agreement with the fact that both Asp-32 and Asp-215 have reacted with EPNP in penicillopepsin crystals at room temperature.

Although it is possible to accommodate two EPNP molecules simultaneously in this active site region of penicillopepsin, it could also be that the binding of one EPNP molecule to either of Asp-32 or Asp-215 would preclude the binding of the second EPNP molecule and that we are actually observing statistical disorder in the crystals. This is unlikely, as the two peaks of electron density (Fig. 2) are of approximately equal size, and the interplanar separation of the *p*-nitrophenoxy moieties is of the order of $3.5 \text{ \AA}$, a reasonable value for aromatic rings. The residues of penicillopepsin that are close to the EPNP molecule bound to Asp-215 are those residues that are likely to be responsible for binding natural peptide substrates in the S_1 site of penicillopepsin. These residues are Leu-222, Thr-218, Ile-301 and a segment of the hairpin loop from Asn-290 to Ile-297 (see Fig. 1).

The conformational change observed for Tyr-75 is of particular interest because of the analogy of this movement with that observed in substrate binding with carboxypeptidase A_a (refs 11, 12). Tyr-75 is located on a hairpin loop (β -bend) which involves residues 70–83 (see Fig. 1). In the penicillopepsin electron density map, the density corresponding to this 'flap' is very well defined because this region represents a close intermolecular contact between two penicillopepsin molecules across a twofold symmetry axis in the crystals. There is a third strand which anchors the flap into a three stranded antiparallel β -sheet. These strands do not form part of the major 18-stranded β -sheet, but they are made up of stretches of β -chain conformation from residues, 70–77; 78–83; 105–108. Although the conformational changes induced on EPNP binding involve only Tyr-75 (rotation around the C_α – C_β bond), it is probable that this whole 'flap' region undergoes a major conformational change to accommodate the binding of longer substrate molecules. This would be the reason for the CD spectral changes observed when binding certain activators to penicillopepsin¹³.

Fig. 2 Stereoscopic photograph of the difference electron density map of the covalent inhibitor EPNP (see text). Positive values of difference density are represented as solid contours, negative values with dashed contours. The two EPNP molecules bind to the carboxyl groups of Asp-32 and Asp-215 and the *p*-nitrophenoxy moieties are coplanar and point towards the bottom of the photograph. Since this is a difference map the electron density of the carboxyl groups is not seen, but lies approximately where the side chains are designated (see Fig. 3a). The conformational movement of Tyr-75 is seen to the right of and below the positive density of the EPNP molecule bonded to Asp-32.



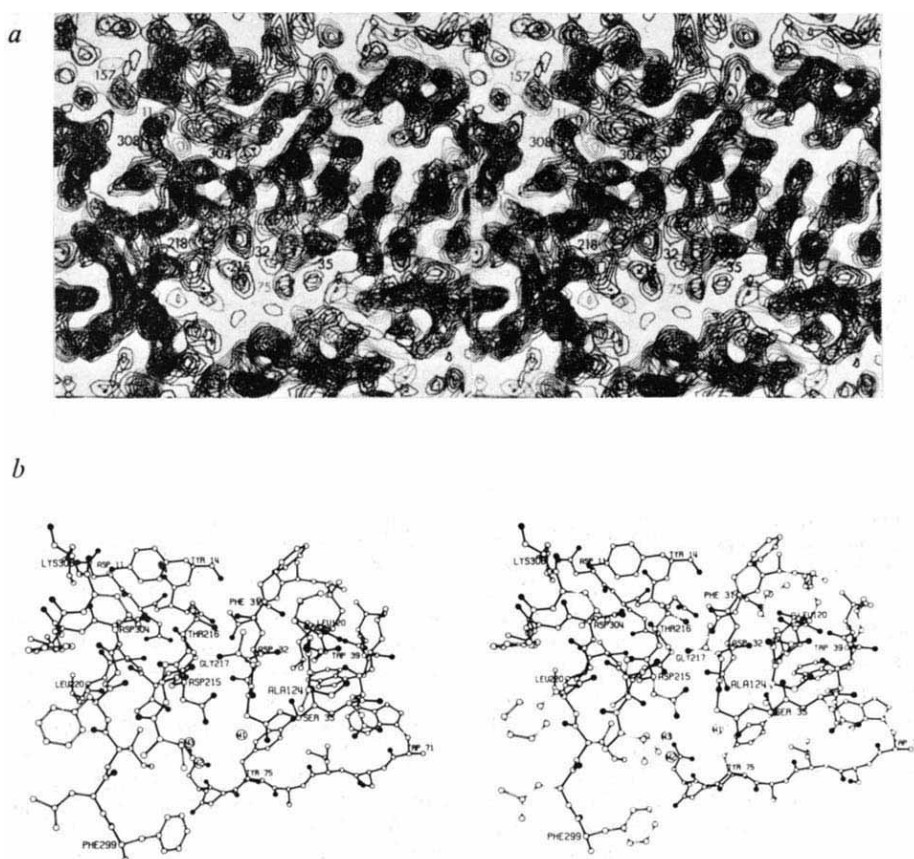


Fig. 3 *a*, Electron density map of penicillopepsin in the region of the active site. Side chains of the amino acids discussed in text have been labelled with the numbering given in ref 1. *b*, Schematic interpretation of the electron density map shown in 3*a* showing the disposition and hydrogen-bond interactions between the several groups in the region of the active site (portions with black connectivity in Fig. 1*b*). The bound water molecules are labelled W(1) W(2) and W(3).

A final point regarding the EPNP studies concerns the difference of the nature of the covalent penicillopepsin-EPNP complexes with Asp-32 and Asp-215. The mechanism of ring opening of 1,2-epoxides in acid solution favours products in which the nucleophile attacks the 2 position of the propane moiety¹⁴. The electron density associated with the EPNP molecule bonded to the carboxyl group of Asp-215 is consistent with this mechanism as there is difference density on either side of the carboxyl group. But, the difference electron density of the EPNP molecule which has reacted with Asp-32 clearly terminates at that group on the enzyme and the complex is Asp-32 (EPNP) with the carboxyl group bonded to the 1 position of the propane moiety. This fact indicates that the carboxyl group of Asp-32 is inaccessible for the formation of covalent intermediates with peptide substrates without major conformational changes in the enzyme's tertiary structure. A negative peak on the difference map in the region between the two carboxyl groups of Asp-32 and Asp-215 shows that the hydrogen-bonded complex of these two groups in the native enzyme structure has been disrupted by their reaction with EPNP.

Disposition of groups in the active site of penicillopepsin

Before considering the catalytic mechanism of acid proteases, we feel it is important to emphasise further that the two carboxyl groups located in the active site are relatively inaccessible to substrate molecules, and that most of the region is inflexible to gross conformational changes. A carboxylate anion (Asp-32) and a protonated carboxyl group (Asp-215) have been implicated in the catalytic mechanism of acid proteases (ref. 3 and references cited therein). The sections of the electron density map which contain these groups are shown in Fig. 3*a*. Figure 3*b* is a schematic representation of the corresponding part of the molecule. It can be seen that the carboxyl group of Asp-32 is in an unusually tight hydrogen-bonded environment reminiscent of the hydrogen-bonded environment of Asp-102 of the serine

proteases¹⁵. The contacts that are made to the carboxyl group of Asp-32 include (1) the carboxyl group of Asp-215, (2) the O^γ of Ser-35, (3) the main chain N-H of Gly-217 and (4) a bound water molecule W(1). Although not seen in Fig. 3, the main chain from Asp-32 to Gly-34 traverses directly above (towards the viewer) the carboxyl group of Asp-32 and makes several close contacts with this moiety.

The carboxyl group of Asp-215 is more exposed to solvent than is the carboxyl group of Asp-32. The side chain of Thr-218 is located to the left and below (away from the viewer in Fig. 3*a*) the COOH group of Asp-215 but no direct hydrogen bonding exists between these two side chains. Two water molecules, W(2) and W(3), are within hydrogen-bonding distance of Asp-215. The side chain Ile-213 is seen above and to the left of Asp-215 and in conjunction with the side chains of Phe-189 and Tyr-299 probably forms the S₁' hydrophobic binding site in penicillopepsin¹⁶. These residues correspond to Ile-213, Tyr-189 and Leu-299 respectively, in pepsin¹⁷.

There are other residues in the neighbourhood of the two active site aspartic acid groups. The peptide bond from Thr-216 to Gly-217 lies between the two carboxyl groups of Asp-32 and Asp-304. The carboxyl group of Asp-304 (Fig. 3), deeply buried in the hydrophobic core of the molecule, makes a hydrogen-bonded contact with the main chain C=O of Thr-216. To the left of Asp-304 there is an ion-pair, formed by Lys-308 and Asp-11, which has neighbouring residues Tyr-14 and His-157 to its right and left respectively. Tyr-14 is also buried in the hydrophobic core and the hydroxyl group lies within 3.2 Å of Asp-304 and the ion-pair Lys-308 to Asp-11. All of the above mentioned residues with the exception of His-157 are conserved in the porcine pepsin sequence^{1,17}.

In addition to these side chain interactions it is important to consider the polypeptide chain conformation in the vicinity of the active site aspartic acid residues. Both residues are in regions of chain that show a high degree of not only primary structural homology but also tertiary structural homology¹. The chain from Phe-31 to Ser-35 adopts a similar conformation

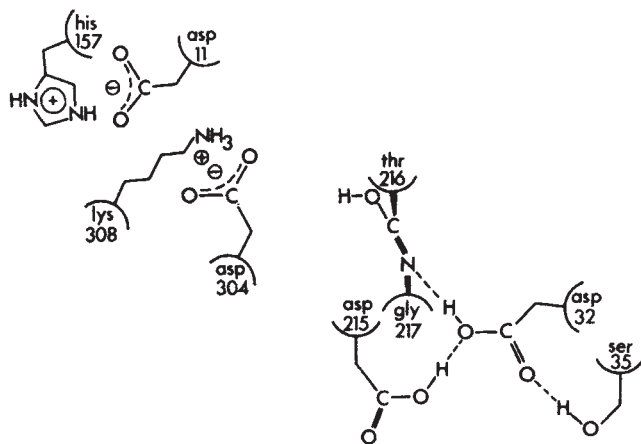


Fig. 4 Proposed tautomeric form of the peptide bond between Thr-216 and Gly-217 which could result from conformational changes induced on the binding of extended substrate peptides to penicillopepsin. The net effect is to enhance the electrophilic character of the proton between Asp-215 and Asp-32.

to the segment of chain from Ala-214 to Thr-218. (There is in fact much more extensive tertiary structural homology between the two lobes of penicillopepsin, but it is not pertinent to the discussion here and will be considered in a subsequent publication). Asp-32 is in strand c (strand nomenclature for penicillopepsin and the corresponding residue numbers within strands are given in ref. 1), with nearest neighbours strand h (running parallel to c) and strand m (anti-parallel to c) which contains Asp-215. In addition to strand c, strand m (containing Asp-215) has a parallel strand p which is its other nearest neighbour strand. These main chain conformations and the fact that the strands are in an extensive mixed β sheet which has a pronounced hydrophobic side, lead one to the inference that the active site region is conformationally rigid and that movement of the two lobes relative to one another would not make the active site much more accessible to substrate. In support of this, it has been observed that pepsin and penicillopepsin retain partial catalytic activity in 8 M urea (T. Hofmann, unpublished).

The immediate surroundings of the two active site carboxyl groups are consistent with the observations that Asp-32 has a low pK_a and Asp-215 has a high pK_a , and the latter would be predominantly protonated at pH values from 2 to 4. As this proton is shared between these two carboxyl groups, however, there would be little driving force for it to act in the role of facilitated proton transfer to the substrate carbonyl oxygen atom in the manner suggested by Knowles¹⁸. The presence of the carboxyl group of Asp-304 adjacent to the peptide bond between Thr-216 and Gly-217 offers a charge relay system which is the major factor giving this shared proton an electrophilic character. Enhanced electrophilicity of this proton may be achieved by the mechanism depicted schematically in Fig. 4. Asp-11 is at the extremity of a hairpin loop (strands a and b, ref. 1) in the central portion of the large mixed β -sheet region that would be involved with the binding of an extended substrate molecule. A relatively minor conformational change of this loop would disrupt the Asp-11 to Lys-308 ion pair, and at least in penicillopepsin, Asp-11 could form an alternative ion pair with His-157 which is on the surface of the enzyme and is protonated at pH values < 5 . This induced conformational change would leave an almost completely buried positive charge of Lys-308 in the hydrophobic interior. This energetically unfavourable situation could be relieved by a subsequent conformational change in which the carboxyl-group of Asp-304 protonates the main chain carbonyl oxygen atom of Thr-216 and rotates by $\sim 120^\circ$ about the $C_\alpha-C_\beta$ bond to form an internally buried ion-pair with Lys-308. The protonation of the peptide bond between Thr-216 and Gly-217 would then produce

the tautomer as shown in the diagrammatic sketch in Fig. 4. Support for this proposal comes also from chemical studies of Kitson and Knowles¹⁹ and Huang and Tang²⁰ in which it was observed that chemical modification of arginine residues of pepsin resulted in a loss of catalytic activity. Chemical modification of Lys residues of penicillopepsin results in a similar loss of activity (T. Hofmann, unpublished). As the only basic residue that is in common between porcine pepsin and penicillopepsin¹ is at position 308 we suggest that Arg-308 is the modified residue in porcine pepsin.

These conformational changes involving Asp-11, Lys-308 and Asp-304 represent an extreme case in which a very strong acid, the protonated peptide bond between Thr-216 and Gly-217, is formed. We do not suggest that this extreme conformational change is ever reached, however. A partial contribution towards the tautomer shown in Fig. 4 would also result in enhanced electrophilic character of the proton which is bound between the two carboxyl groups Asp-32 and Asp-215, rendering this proton more labile than one would expect in its shared state as seen in the penicillopepsin structure. The degree of electrophilic character of the proton would then be a function of the relative contributions of these two tautomeric forms of the Thr-216 to Gly-217 peptide bond.

Our structural studies have shown that penicillopepsin has an extended substrate binding groove¹. A protein conformational change is known to occur when certain tripeptide activators (Leu-Gly-Leu) bind in a region near Trp-39, Trp-71 and Tyr-75 (ref. 13). Transpeptidation products isolated from reaction mixtures in which the activators are present indicate that this binding region (secondary A site as defined by Takahashi and Hofmann¹⁶) is the N-terminal region of bound substrate, that is, P_1-P_4 . In addition, Fruton has shown that increased substrate chain length in the N-terminal direction from the scissile bond enhances k_{cat} with little difference in substrate binding affinity³. These effects may now be explained by the increased electrophilic character of the proton between the two carboxyl groups of Asp-32 and Asp-215.

Towards a catalytic mechanism for acid proteases

The acyl-enzyme mechanism for serine proteases is well established¹⁵ and the postulated tetrahedral intermediate in that pathway has subsequently been observed in crystalline trypsin-trypsin inhibitor complexes^{21,22}. Catalytic mechanisms for the thiol proteases²³ and metalloproteases¹² have also been advanced and the proponents of these mechanisms have drawn heavily on the results of crystal structural studies. In the most recent review article on acid protease mechanisms³, Fruton states at the outset "that speculations about the catalytic mechanism of pepsin have been limited by lack of information about the three-dimensional structure of its active site . . .". The electron density map of penicillopepsin at 2.8-Å resolution¹ provides the missing details of the active site geometry and the following proposed mechanism is based on the structural results and on chemical and kinetic studies done with synthetic substrates of acid proteases.

The description of the active site geometry of penicillopepsin in the previous section, by no means exhaustive, was intended to convey two major points: (1) that the β -carboxyl group of Asp-32 is too buried for this group to act as an attacking nucleophile directly, and (2) that the sum of the interactions of the groups in the active site and the conformations of the main polypeptide chains prohibit large conformational changes in this region. A conformational change involving the extensive anti-parallel β sheet which contains Tyr-75, as discussed above, is possible however, and probably accounts for the conformational changes observed on binding certain peptide activators. It is also unlikely that the carboxyl group of Asp-215 acts as an attacking nucleophile even though it is partially exposed to solvent, because it is the carboxyl group with the higher apparent pK_a of 4.7 (ref. 18). How then is one to accommodate peptide bond catalysis by acid proteases?

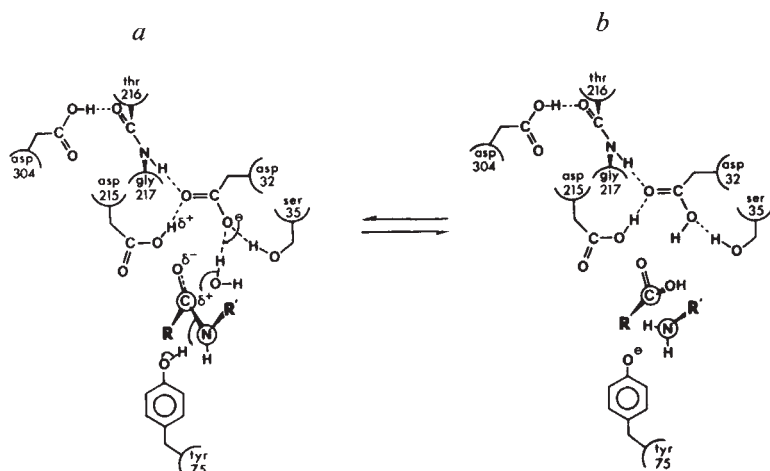


Fig. 5 Proposed general base catalytic mechanism of the acid proteases. *a*, Schematic drawing of the productive binding mode for substrate RCONHR' in the active site of penicillopepsin. The conformational change of the 'flap' region brings Tyr-75 to a position from which it can protonate the amide nitrogen of the scissile bond. Asp-32 attacks the carbon atom of the polarised $C^{6+} \cdots O^{6-} \cdots H^{6+}$ system through the bound water molecule W(1). Transpeptidation through acyl or amino transfer is a function of the relative affinity of the enzyme for the acyl or amino fragment shown in *b*.

The scheme that we propose is shown in Fig. 5 which contains the essential details of the active site (shown more extensively in Fig. 3). It also includes the scissile peptide bond of a hypothetical substrate RCONHR'. The structure of the binding region, as observed in the Kendrew model of penicillopepsin, is compatible with that orientation of an X-Phe-Phe-Y type substrate (the Phe-Phe bond being cleaved). Side chain interactions of substrate with enzyme have been omitted from Fig. 5 to enhance clarity, but are essentially as described in the previous sections.

The formation of the Michaelis complex is accompanied by a conformational change of the flap region (residues 74–80) which has to open outwards to admit the substrate into the extended substrate binding cleft. We have pointed out an approximation to this change in the EPNP binding experiment, where it was limited to the movement of Tyr-75 by the crystal packing of molecules of penicillopepsin. Substrate binding also displaces water molecules W(2) and W(3), whereas W(1) can remain in its original position. It now lies between the substrate (as postulated) and the carboxyl group of Asp-32. With the substrate bound, the flap can return and in so doing, the phenolic OH of Tyr-75 comes close to the amide of the susceptible peptide bond. The productive binding mode is now complete as depicted in Fig. 5*a*.

It is useful to point out the analogy of this binding mode to the substrate orientation found for glycyltyrosine in the carboxypeptidase A study done by Lipscomb and co-workers^{11,12}. In our case, Tyr-75 is in a position to protonate the amide nitrogen atom of the peptide bond (Fig. 5*a*). This function is carried out by Tyr-248 in carboxypeptidase A after a similar conformational change of the phenolic group upon substrate binding. The electrophilic component of the carboxypeptidase mechanism is ascribed to the Lewis acid, Zn^{2+} in the Zn-carbonyl mechanism. Polarisation of the carbonyl bond in the case of acid proteases is accomplished by the proton shared between the two active site carboxyl groups. We propose that the electrophilicity of this proton is a function of the nature of the substrate, and could be enhanced by specific enzyme-substrate interactions as described in the previous section.

Finally, the Zn-carbonyl mechanism for carboxypeptidase A postulates either a direct nucleophilic attack of Glu-270 on the polarised carbonyl carbon of the peptide bond or a general base attack through a bound water molecule. The analogous residue to Glu-270 of carboxypeptidase A is the carboxylate of Asp-32 of the acid proteases. Carboxypeptidase A has a *pH* optimum close to 7.5 where the Glu-270 side chain would have a carboxylate anion. The special requirement of acid proteases in having an ionised carboxylate at very low *pH* values are met by the most unusual environment of the Asp-32 side chain. We propose that attack of the polarised carbonyl carbon atom of the scissile peptide bond occurs through the activated water molecule W(1) rather than by direct nucleophilic attack by the

β -COO⁻ of Asp-32 (Fig. 5*a*). Thus we prefer this general base mechanism of hydrolysis without the involvement of either an anhydride or an amino-enzyme intermediate.

A major stumbling block in the various other proposed mechanisms of acid proteases, as reviewed by Fruton³, has been the inability of workers to trap either of the covalent intermediates (acyl-enzyme or amino-enzyme). Our proposal obviates this problem and postulates a ternary complex of enzyme, acyl-fragment and amino-fragment after cleavage of the susceptible peptide bond. The order of release of products is determined by the relative affinity of the acyl or amino moiety to their respective binding sites²⁴.

Transpeptidation is a phenomenon peculiar to the acid proteases and it has only been observed to proceed to any significant degree at *pH* values greater than 4.0. Our mechanism is also compatible with the transpeptidation phenomenon. The products resulting from hydrolysis of the peptide bond are most likely the carboxylate ($R-COO^-$) and the protonated amino group (NH_3^+-R'). Transfer of the proton from the carboxyl group to the amino group would occur readily and the positively charged amino group could be stabilised by the phenoxide ion of Tyr-75 (Fig. 5*b*). As an example, consider the work of Newmark and Knowles²⁴ on pepsin, in which they studied the transpeptidation products from the peptide ^{14}C -Leu-Tyr- 3H -Leu. Approximately 80% of the resulting Leu-Leu is ^{14}C -Leu- ^{14}C -Leu; the remainder is 3H -Leu- 3H -Leu with none of the mixed radiolabelled product. This tripeptide binds initially such that the N-terminal Leu-Tyr bond is cleaved. Exchange of Tyr-Leu with a second Leu-Tyr-Leu is more facile than the exchange of the bound N-terminal ^{14}C -Leu product from the first hydrolysis. Transpeptidation would occur through the microscopic reversibility of the mechanism shown in Fig. 5, yielding the Leu-Leu-Tyr-Leu which is subsequently cleaved by a second hydrolytic step to give ^{14}C -Leu- ^{14}C -Leu. Fruton³ has pointed out that at *pH* values at which transpeptidation occurs (>4.0) the equilibrium constant for the reaction: $RCOO^- + H_3N^+-R' \rightleftharpoons RCONHR' + H_2O$ is 1 and reversal of the hydrolytic step is equally probable. But, at *pH* values close to the optimum *pH* of pepsin (*pH* ~ 2) one would expect the carboxyl of the acyl fragment to be protonated, thus precluding transpeptidation at low *pH* values.

Conclusions

The mechanism proposed here is consistent with the disposition of amino-acid groups at the active site of penicillopepsin and with various characteristics of acid protease hydrolysis. The electrophilic component of hydrolysis is the shared proton between Asp-32 and Asp-215 (ref. 18). The electrophilicity of this proton is a function of specific enzyme-substrate interactions and its variation is seen in the variation of k_{cat} with closely related substrates. We have not used hypotheses that require covalent acyl or amino intermediates because the

existence of these species has not been detected or conclusively proven. Whether or not transpeptidation proceeds through acyl or amino transfer (indeed both pathways have been observed in the same reaction mixture) is determined by the affinity of the amino or acyl moiety bound to the enzyme following the first hydrolytic step. Transpeptidation proceeds only at pH values where one has charged groups involved with the formation of the new peptide bond. Protonation of the carboxylate at low pH values precludes transpeptidation through this mechanism and favours hydrolysis.

Finally, another convincing argument in favour of this mechanism is the very close similarity to that one proposed by Lipscomb for carboxypeptidase A. It seems that these two enzymes have converged on a mechanism for the hydrolysis of peptide bonds which differs from that of the serine protease mechanism. The specific requirement of catalytic activity in strongly acid media precludes the use of a Lewis acid such as Zn^{2+} . The acid proteases have instead developed a bond-polarising system with a bound proton and a charge-relay system which influences the electrophilicity of this proton. Tyr-75 clearly is very similar to Tyr-248 of carboxypeptidase both in its function of transferring a proton to the amide nitrogen atom and the conformational change it undergoes on substrate binding. Both the metalloproteases^{12,25} and the acid proteases use an ionised carboxylate group to catalyse the conversion of the substrate carbonyl carbon to a transient tetrahedral intermediate which results in the cleaved peptide bond. The proposal here that acid proteases catalyse this step through general base attack through a bound water molecule (W(1)) lends support to this step being present as well in the carboxypeptidase A mechanism. It has been pointed out that the carboxypeptidase A mechanism resembles that of pepsin in several respects²⁶ and it is therefore gratifying to find that the structural features of penicillopepsin are consistent with such closely corresponding mechanisms of catalysis.

We thank Professors Jeremy Knowles and Theo Hofmann for helpful suggestions. This work was supported by MRC of Canada.

Note added in Proof: Completion of the building of the entire model of penicillopepsin has required the following revisions to the sequence reported in ref 1. (a) The molecule is made up of 323 residues. (b) The following residues are now interpreted as: 299 is Phe, 302 is Phe and 249–264 are tentatively assigned as: -Asx-Cys-Ser-Ser-Ser-Val-Pro-Asx-Phe-Ser-Val-Ser-Ile-Ser-Gly-Tyr-.

Received 17 March; accepted 10 May 1977.

- 1 Hsu, I.-N., Delbaere, L. T. J., James, M. N. G. & Hofmann, T. *Nature* **266**, 140–145 (1977).
- 2 Subramanian, E. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 556–559 (1977).
- 3 Fruton, J. S. *Adv. Enzym.* **44**, 1–36 (1976).
- 4 Hsu, I.-N., Hofmann, T., Nyburg, S. C. & James, M. N. G. *Biochem. biophys. Res. Commun.* **72**, 363–368 (1976).
- 5 Coddling, P. W., Delbaere, L. T. J., Hayakawa, K., Hutcheon, W. L. B., James, M. N. G. & Jurasek, L. *Can. J. Biochem.* **52**, 208–220 (1974).
- 6 Delbaere, L. T. J., Hutcheon, W. L. B., James, M. N. G. & Thiessen, W. E. *Nature* **257**, 758–763 (1975).
- 7 Blow, D. M. & Crick, F. H. C. *Acta Crystallogr.* **12**, 734–802 (1959).
- 8 Cunningham, A. *et al. Can. J. Biochem.* **54**, 902–914 (1976).
- 9 Hartsuck, J. A. & Tang, J. J. *biol. Chem.* **247**, 2575–2580 (1972).
- 10 Chen, K. C. S. & Tang, J. J. *biol. Chem.* **247**, 2566–2574 (1972).
- 11 Lipscomb, W. N. *et al. Brookhaven Symp. Biol.* **21**, 24–90 (1968).
- 12 Quirocho, F. A. & Lipscomb, W. N. *Adv. Protein Chem.* **25**, 1–78 (1971).
- 13 Wang, T. T., Dorrington, K. J. & Hofmann, T. *Biochem. biophys. Res. Commun.* **57**, 865–869 (1974).
- 14 Parker, R. E. & Isaacs, N. S. *Chem. Rev.* **59**, 737–799 (1959).
- 15 Blow, D. M. *Acc. Chem. Res.* **9**, 145–152 (1976).
- 16 Takahashi, M. & Hofmann, T. *Biochem. J.* **147**, 549–563 (1975).
- 17 Sepulveda, P., Marciniuszyn, J., Liu, D. & Tang, J. J. *biol. Chem.* **250**, 5082–5088 (1975).
- 18 Knowles, J. R. *Phil. Trans. R. Soc. B257*, 135–146 (1970).
- 19 Kitson, T. M. & Knowles, J. R. *FEBS Lett.* **16**, 337–338 (1971).
- 20 Huang, W.-Y. & Tang, J. J. *Fedn Proc.* **30**, 759 (1971).
- 21 Rühlmann, A., Kukla, D., Schwager, P., Bartels, K. & Huber, R. *J. molec. Biol.* **77**, 417–426 (1973).
- 22 Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H. & Blow, D. M. *Biochemistry* **13**, 4212–4228 (1974).
- 23 Dreath, J., Jansonius, J. N., Koekoek, R. & Wolthers, B. G. *Adv. Protein Chem.* **25**, 79–115 (1971).
- 24 Newmark, A. K. & Knowles, J. R. *J. Am. chem. Soc.* **97**, 3557–3559 (1975).
- 25 Kaiser, E. T. & Kaiser, B. L. *Accts Chem. Res.* **5**, 219–224 (1972).
- 26 Fruton, J. S. *Accts Chem. Res.* **7**, 241–246 (1974).

letters to nature

Intensity transitions in Cyg XR-1 observed at high energies from OSO 8

THE high energy X-ray source, Cygnus XR-1, exhibits a variable intensity between states which differ in flux at lower energies (typically 3 to 6 keV) by a factor of three (refs 1–5). Although the transitions between these 'high' and 'low' flux states occur at irregular intervals, the transition itself usually takes no more than one day in either direction. We report here an observation of such a transition at energies above 20 keV which shows that the spectrum of Cyg XR-1 exhibits the pivoting effect during intensity transitions expected from two-temperature accretion disk models of the X-ray emitting region^{6,7}.

Cyg XR-1 was observed with the high energy X-ray spectrometer on board the OSO-8 satellite from 11 to 19 November, 1975, and from 27 October to 15 November, 1976 (excluding the period from 1 November to 7 November, 1976). The CsI(Na) scintillation spectrometer is sensitive to radiation in the 20 keV to 3 MeV range. The detector is actively shielded and collimated to a 5-degree circular FWHM and offset by an angle of 5 degrees from the negative spin axis of the satellite. Further details of the instrument and its satellite environment are given by Dennis *et al.*⁸ Once every rotation period of the satellite (~10 s), any source on the detector's scanning circle will transit through its field of view. The data were analysed in the manner outlined by Dolan *et al.*⁹.

The raw counting rates during the 1975 observations in the interval 23 to 153 keV are shown in Fig. 1. Data are grouped in intervals of 0.55995 d, or a phase interval of one-tenth the period of the HDE 226868 binary system¹⁰. A decrease in counting rate of ~40% was observed to occur on 16 November, 1975, with the source remaining at a significantly lower counting rate after the transition than before it. Comparison with 3–6-keV proportional counter data taken by Holt *et al.*¹ and Serlemitsos (quoted in ref. 11) reveals a simultaneous rise in the lower energy counting rate. Boldt¹² distinguishes between 'transitions' and 'transient events'. In transient events, the 3–6-keV intensity fluctuates by a large amount from its previous value but returns to its original level within a few binary periods; in transitions, the flux remains at its new level for many binary periods. We cannot distinguish between these two possibilities with our observations, and shall refer to all large variations lasting more than a few days as transitions. It is apparent that we have observed for the second time at higher energies¹³ a transition between intensity states in Cyg XR-1.

Our observations for 1976 are shown in Fig. 2. The 23–153-keV intensity was higher than that observed in 1975, indicating that a reverse transition must have occurred between the two observations, in agreement with the report of such a transition in February 1976, by Holt *et al.*¹ The decrease observed by us to occur on 14 November 1976, probably does not correspond to a change in intensity state at low energies;

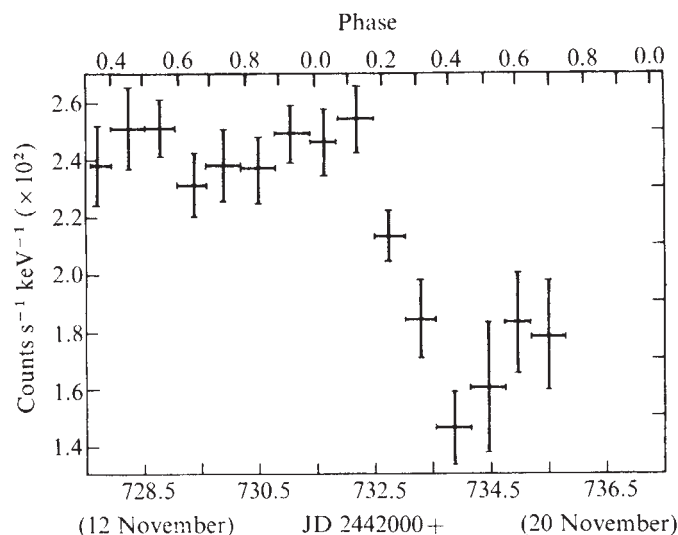


Fig. 1 The 23–153-keV counting rates observed from Cyg XR-1 in November 1975.

the 23–153-keV flux after the transition is still above that observed by us before the transition in 1975. Instead, it illustrates the large intrinsic variability present in the source even when in the ‘low’ state at lower X-ray energies. No 3 to 6 keV proportional counter data are available for comparison at the time of writing.

Typical spectra taken during periods of high and low intensity are shown in Fig. 3. The higher intensity spectrum, observed between phases 0.50 and 0.60 starting JD 2443092.505 (10 November 1976) is well fitted by a power law spectrum $dN/dE = C(E/E_0)^{-\alpha}$ photon $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$ where $C = (1.8 \pm 0.1) \times 10^{-3}$ photon $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$, $\alpha = 2.1 \pm 0.1$ and $E_0 = 73.6 \text{ keV}$; the lower intensity spectrum, observed between phases 0.40 and 0.50 starting JD 2442733.583 (17 November 1975), is well fitted by a power law spectrum with $C = (1.0 \pm 0.2) \times 10^{-3}$ photon $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$, $\alpha = 2.4 \pm 0.4$ and $E_0 = 57.2 \text{ keV}$. The quoted uncertainties are 68% confidence intervals calculated according to the method outlined by Lampton *et al.*¹⁴. The anti-correlation between intensities at lower and higher energies indicates that the spectrum must pivot about a point between the two energy ranges. From the observations of Coe *et al.*¹³, this pivot point occurs at $\sim 7 \text{ keV}$.

All individual spectra that we observe, measured over 0.6-d intervals, can be well fitted by a power law, although the

constant C varies from spectrum to spectrum. The resultant uncertainties in α obtained from fits in our limited energy range do not allow us to show a detailed correlation between observed spectral index and the 23–153-keV counting rate. All the 1976 data are at higher intensity than the 1975 data, however, and the 1976 data have a lower average value of α (2.05 ± 0.03) than the 1975 data ($\alpha = 2.39 \pm 0.04$), as would be expected from the discussion above. If these average spectra extend unchanged to lower energies, the 2–300-keV luminosity of the system is in fact greater when the source is more intense at lower energies (that is, the ‘high’ state at 3–6-keV is indeed the higher luminosity state of the system).

This pivoting behaviour (or ‘see-saw’ effect) is predicted to occur in models for which the X-ray emission is produced in a two-temperature accretion disk surrounding the secondary component of the binary system^{6,7}. In these models, the inner part of the accretion disk, ballooned up by radiation pressure, is geometrically thick but optically thin to higher energy X-rays. The high-energy radiation, predicted to have a power law spectrum, is produced from the lower energy X-ray or ultraviolet photons by inverse Compton scattering off relativistic electrons in this region of the disk. Basically, an increase in the lower energy photon flux ‘cools off’ the electrons by causing them to undergo more inverse Compton collisions, leading to a lower mean energy gained by the photons per scattering and, ultimately, to a lower flux of high energy photons. This anti-correlation between the fluxes in the two energy ranges would then be observed as the same type of pivoting of the spectrum that we detect in Cyg XR-1.

Transitions in Cyg XR-1 are observed to occur at irregular intervals, but seem to occupy much the same duration, about one day, regardless of direction of the transition or interval since the last transition^{1–5}. We note that not only the two transitions we observed (one of which may not be associated with a large change in the low energy flux) but also six of the seven other transitions in the intensity state of Cyg XR-1 reported in refs 1–5 commenced near periastron at phase 0.18 (ref. 10). It is tempting to speculate that transitions in the luminosity are linked to variations in the mass transfer rate from the primary to the accretion disk surrounding the secondary. Alme and Wilson¹⁵ have proposed a model of the system in which matter is transferred primarily by radiation-driven density waves generated in the atmosphere of the primary, which fills only about 85% of its Roche lobe. In such a model having the orbit eccentricity and period of the HDE 226868 system, nearly all the mass is transferred in an interval approximately 1 d long near periastron. As pointed out by Alme and Wilson, the period between mass transfer ‘pulses’

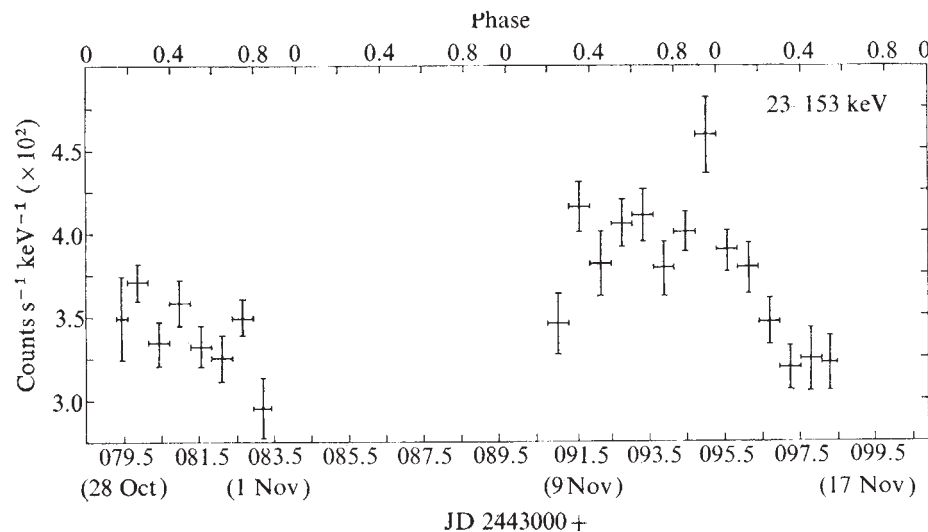


Fig. 2 The 23–153-keV counting rates observed from Cyg XR-1 in October and November 1976. Note the difference in the scale of the ordinate from Fig. 1. Spacecraft pointing requirements prevented observations between 1 and 7 November.

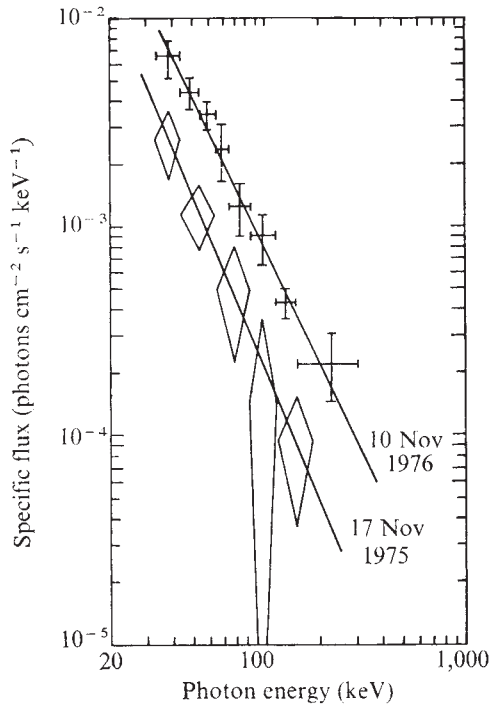


Fig. 3 Typical spectra taken during a 3–6-keV 'low' state (crosses, above), and a 'high' state (diamonds, below). The straight lines are power law fits to the spectra; spectral parameters are given in the text. Error bars are ± 1 standard deviation in the flux. Individual channels in the original data have been grouped to improve statistics.

is not critically phase locked and the period between pulses is small compared with the characteristic disk transfer time. Hence, variations in the X-ray emission at this period would be difficult to detect if approximately the same amount of mass were transferred every pulse. If the primary, for some unspecified reason, suddenly changed drastically the amount of matter it transferred per pulse, however, the disk might then be expected to show a changed luminosity commencing near periastron.

A large increase in mass transferred per pulse would presumably increase the low energy photon flux in the outer regions of the disk. The high energy flux would then follow in an anti-correlated fashion, as discussed above. The difference between 'transitions' and 'transient events'¹² would then be related to whether the mass transferred in succeeding pulses was maintained at its new value or returned to its previous level. It will be interesting to see whether this correlation of commencement time with phase near periastron is borne out in future transitions.

J. F. DOLAN
C. J. CRANNELL
B. R. DENNIS
K. J. FROST
L. E. ORWIG

NASA-Goddard Space Flight Center,
Greenbelt, Maryland 20771

Received 11 April; accepted 4 May 1977.

- ¹ Holt, S. S., Kaluzienski, L. J., Boldt, E. A. & Serlemitsos, P. J. *Nature* **261**, 213 (1976).
- ² Tananbaum, H., Gursky, H., Kellogg, E., Giacconi, R. & Jones, C. *Astrophys. J.* **177**, L5 (1972).
- ³ Kaluzienski, L. J., Holt, S. S., Boldt, E. A. & Serlemitsos, P. J. *IAU Circ.* No. 2863 (1975).
- ⁴ Walker, E. N., Brownlie, G. D., Mason, K. O., Sanford, P. W. & Quintanilla, A. R. *Nature* **263**, 393 (1976).
- ⁵ Sanford, P. W., Ives, J. C., Bell Burnell, S. J., Mason, K. O. & Murdin, P. *Nature* **256**, 109 (1975).
- ⁶ Shapiro, S. L., Lightman, A. P. & Eardley, D. M. *Astrophys. J.* **204**, 187 (1976).
- ⁷ Eardley, D. M. & Lightman, A. P. *Nature* **262**, 196 (1976).
- ⁸ Dennis, B. R., Frost, K. J., Lencho, R. J. & Orwig, L. E. *Space Sci. Instr.* (in the press).

- ⁹ Dolan, J. F., Crannell, C. J., Dennis, B. R., Orwig, L. E., Maurer, G. S. & Frost, K. J. *Astrophys.* (in the press).
- ¹⁰ Bolton, C. T. *Astrophys. J.* **200**, 269 (1975).
- ¹¹ Dolan, J. F., Dennis, B. R., Crannell, C. J., Frost, K. J. & Orwig, L. E. *Highlights of Astronomy* **4** (in the press).
- ¹² Boldt, E. *X-Ray Binaries*, NASA SP-389, 369 (1976).
- ¹³ Coe, M. J., Engel, A. R. & Quenby, J. J. *Nature* **259**, 544 (1976).
- ¹⁴ Lampton, M., Margon, B. & Bowyer, S. *Astrophys. J.* **208**, 177 (1976).
- ¹⁵ Alme, M. L. & Wilson, J. R. *Astrophys. J.* **210**, 233 (1976).

On the association between a γ -ray burst and a radio pulse

ONE minute after the γ -ray burst of 16 August 1976, a radio event with ~ 30 s duration was recorded by three different instruments, a dipole array and a 10-m parabola at Medicina (Bologna), and a 10-m parabola at Trieste (400 km away) (refs 1 and 2). The peak intensity at 151 and 408 MHz was $\sim 10^{-21}$ W m $^{-2}$ Hz with an upper limit to the dispersion between the two frequencies of ~ 10 s. In the absence of specific data on the γ -ray flux we take a typical value of 10^{-4} erg cm $^{-2}$ corresponding to an energy output of $10^{38} d_{100}^2$ erg, where d_{100} is the distance in units of 100 pc.

Here we shall assume that the γ - and radio pulses are physically associated and discuss some implications on the nature of the source. The duration and time lag of the radio pulse are reminiscent of flares in ultraviolet Ceti stars³ where the main activity is in the optical band. However the larger power and temperature required for γ -ray bursts favour models involving white dwarfs or neutron stars.

The brightness temperature of the radio pulse is

$$T_b = (3.2 \times 10^{41})/R^2 d_{100}^2 \text{ K} \quad (1)$$

for a source dimension $R \simeq 10^6$ cm and $d_{100} \simeq 1$, $T_b \simeq 10^{23}$ K —of the same order as found in pulsars. This implies a coherent emission mechanism. Assuming that this is synchrotron radiation, a number of constraints leading to an estimate of the relevant physical parameters can be arrived at. The synchrotron power is⁵ given by the expression

$$\frac{1e^2c^2}{2\pi c \rho^2} \gamma^4 n V N = 4.5 \times 10^{32} \alpha d_{100}^2 \text{ erg s}^{-1} \quad (2)$$

where the observed radio power has been introduced on the right side, α is the beaming factor ($\alpha = 1$ if the emission is isotropic), n is the density of relativistic particles with Lorentz factor γ , ρ is the curvature radius of the particle trajectories, V is the volume of the emitting region and N the coherence factor, which must satisfy the relation

$$N \geq KT_b/\gamma m_e c^2 \quad (3)$$

If the observation frequency is the critical synchrotron frequency one has

$$v_{cr} = \gamma^3 \frac{c}{2\pi\rho} = 4 \times 10^8 \text{ Hz} \quad (4)$$

Minimising the energy requirements we further impose that the relativistic particle energy is radiated in the radio band during the crossing time of the emission region R/c

$$\gamma m_e c^2 n V = 4.5 \times 10^{32} R/c \alpha d_{100}^2 \text{ erg} \quad (5)$$

Equations (1), (2), (3), (4) and (5) yield

$$N = 3 \times 10^{15} \gamma^3/R \quad (6)$$

$$N \geq 5.5 \times 10^{31} d_{100}^2/(\gamma R^2) \quad (7)$$

For ordinary synchrotron radiation in a dipole field an addi

tional relation is

$$\rho = \frac{\gamma m_e c}{3 \times 10^8 \theta B_{s6}} \left(\frac{R_g}{R_{s9}} \right)^3 \quad (8)$$

where $\gamma m_e c$ is the electron momentum in eV/c, θ is the pitch angle, and $B_{s6} = 10^{-6} B_s$ (G) is the surface magnetic field and $R_{s9} = 10^{-9} R_s$ (cm) is the radius of the star.

The system gives

$$R_g \geq 1.36 \times 10^2 \theta^{2/7} R_{s9}^{6/7} B_{s6}^{2/7} d_{100}^{2/7} \quad (9a)$$

$$\gamma \geq 20 \theta^{-1/14} B_{s6}^{-1/14} R_{s9}^{-3/14} d_{100}^{6/14} \quad (9b)$$

$$n \leq 10^4 \alpha \theta^{-0.5} B^{-0.5} R_g^{3/14} d_{100}^{10/7} \text{ cm}^{-3} \quad (9c)$$

In the limit where $\theta \rightarrow 0$, the radiation due to particles following the field lines reduces to 'curvature' radiation. For $\rho \simeq R$, one has

$$R_g \geq 4 \times 10^{-2} d_{100}^2 \quad (10a)$$

$$\gamma \geq 140 d_{100}^{2/3} \quad (10b)$$

$$n \leq 10^{10} \alpha d_{100}^{-8/3} \text{ cm}^{-3} \quad (10c)$$

The main difference between the two sets of relations is that in the former case even for a small pitch angle, $\theta \sim 10^{-3}$, a much larger extension is implied. Since it seems extremely difficult to produce and maintain coherence conditions over large scales we suggest that curvature radiation is responsible for the radio-emission. The associated synchrotron radiation is emitted at much higher frequency where presumably it is not amplified by coherence. This picture and the derived values of the parameters are similar to those proposed by Pacini and Rees for pulsars⁵.

Because of the upper limit on the signal dispersion it is not possible to explain the 1-m delay between the γ -ray and radio events as caused by the increasing transparency of an expanding cloud. The delay could be either intrinsic or due to a beaming of the radio emission and to rotation of the source. The period should be a few minutes and the total duration of the radio emission comparable. This could explain the previous negative observations⁵ because one expects that only for a small fraction of the γ -ray events the associated radio emission will be detectable at Earth.

If the underlying star is a neutron star the hypothesis that the delay is intrinsic is difficult to support since the Alfvén velocity is large. The rotational energy of a neutron star with $\omega \simeq 10^{-2}$ and its magnetic energy for $B \simeq 10^{12}$ G are both $\simeq 10^{41}$ erg. In the burst 10^{-3} of the available energy should be released and in both cases this should be accompanied by a pulsar glitch⁷. However the fractional energy variation required is much larger than the maximum observed period discontinuities in pulsars. A possible way out is the interaction with an external agent (for example, a companion star), which on the other hand, implies a serious difficulty in explaining the short timescale⁴.

Production of γ -ray bursts by magnetic white dwarfs has been proposed by some authors⁸⁻¹⁰. The mechanism should be a convective instability, which can account for the similarities with ordinary flares. In this model regions with magnetic fields $\sim 10^8$ – 10^9 G are required⁹, while the average surface field B_s can be lower and, for $B_s \sim 10^6$ G, the burst energy is comparable with the total magnetic energy. On the other hand, from the known density of magnetic white dwarfs in the sun vicinity, a burst frequency of 1/40 per year per star is required in order to account for the observed γ -ray event rate. Therefore either the density of magnetic white dwarfs has been underestimated by a large factor or one must suppose that a

dynamo mechanism can reconstruct the magnetic field at the expenses of the rotational energy, which for $\omega = 10^{-2}$ is 10^{47} erg.

Note added in proof: In a recent preprint by Sommer *et al.* serious doubts are raised on a direct association between the radio and γ -ray events of 16 August 1976.

L. MARASCHI
A. TREVES

Istituto di Fisica dell'Università and Laboratorio di Fisica Cosmica e Tecnologie Relative del C.N.R. Milan, Italy

Received 21 March; accepted 26 April 1977.

- ¹ Mandolesi, M., Sironi, G. & Sommer, M. *J.A.U. Circ.* No. 3017 (1976).
- ² Mandolesi, M. *et al.* *Nature* **266**, 427 (1977).
- ³ Spangler, S. R. & Moffett, T. J. *Astrophys. J.* **203**, 497 (1976).
- ⁴ Ruderman, M. *Ann. N.Y. Acad. Sci.* **262**, 164 (1975).
- ⁵ Pacini, F. & Rees, M. J. *Nature* **226**, 622 (1970).
- ⁶ Baird, G. A. *et al.* *Astrophys. J. Lett.* **196**, L11 (1975).
- ⁷ Pacini, F. & Ruderman, M. *Nature* **251**, 399 (1974).
- ⁸ Stecker, F. & Frost, K. *Nature Phys. Sci.* **245**, 70 (1973).
- ⁹ Channugam, G. *Astrophys. J. Lett.* **193**, L75 (1974).
- ¹⁰ Mullian, D. J. *Astrophys. J.* **208**, 199 (1976).

⁴⁰Ar–³⁹Ar age of rocks, and the development mode of the Philippine Sea

THE origin and development of the Philippine Sea have been central issues in tectonic studies of a marginal sea: the deep-sea drilling project (DSDP), Leg 31, was primarily intended to resolve the question¹. Unfortunately, at only two of the Leg 31 sites (292 and 296) were microfossils indicating the age of the basement recovered, so the age of the ocean basin, had to be deduced by dating the drilled basement rocks. There seems to be some confusion in the literature² in interpreting the fossil data with regard to the age of the ocean basin. We have dated several DSDP Leg 31 drilled rocks by the ⁴⁰Ar–³⁹Ar step-heating technique (Table 1 and Fig. 1) and re-evaluated the fossil ages and various hypotheses concerning the development of the western Philippine Sea and the Shikoku Basin. We applied two statistical criteria to judge the reliability of ⁴⁰Ar–³⁹Ar ages, (1) the experimental data with MSUM values (Table 1) less than 3.50 and (2) the apparent ages should agree with each other, within 1σ , for more than 65% of ³⁹Ar released. Although such criteria do not prove the reliability of this dating technique, they provide a measure of it³. Of the nine drilling sites, six reached the basement—site 290, 291, 292, 293, 294 and 296—their ages were determined by both the ⁴⁰Ar–³⁹Ar step-heating dating and micro-paleontological methods (Table 1).

Earlier views as to the development of the Philippine Sea suggested that the western Philippine Basin is Mesozoic or even older, and the CBF was an extinct oceanic ridge^{4,5}. On the basis of the fossil ages obtained from Leg 31, Karig *et al.*⁶ proposed that the Philippine Basin was a complex early Tertiary basin developed at least in part behind an pre-existing arc system. Louden⁷ identified ESE–WNW trending magnetic lineations south of the Central Basin Fault as anomalies 18 through 21. Recently, Watts *et al.*⁸ reported ESE–WNW trending lineation north of the CBF, which they suggested as either Mesozoic or Tertiary in age, although preferring the latter. However, there are no Mesozoic basement ages in the western Philippine Sea and no age progression with distance from the CBF, which would be expected if the CBF is an extinct spreading ridge (Fig. 1).

There seems to be considerable difference between the ⁴⁰Ar–³⁹Ar age and microfossil age for site 292. The lowermost clastic rock (chalk: core 39-section 1–88–99) can be identified to belong to P-15 of Blow's planktonic foraminifera zonation scheme⁹ in the late Eocene. If a timescale of Berggren¹⁰, and Berggren and Van Couvering¹¹ is

Table 1 Age determination of drilled basement rock

Sample	Rocks	Total age† (Myr)	Plateau age‡ (Myr)	⁴⁰ Ar- ³⁹ Ar ages*	MSUM = $\left(\frac{\text{SUMS}}{n-1}\right)^{1/2}$	Best age for the ocean basin (Myr)
				Isochron ages (Myr) and J-value		
DSDP 292-41-2	coarse grained basalt	51.1	49.4 ± 1.5 (for 95% ³⁹ Ar	49.5 ± 2.0 (J = 0.006955)	1.208 for all fractions but 700 °C	49.4
DSDP 293-21-2	gabbro, cumulate rock	440.7	42.0 (73%)	42.0 (J = 0.006955)	—————	42.0
DSDP 294-7-1	alkali basalt	33.9	49.1 ± 3.5 (95%)	48.8 ± 2.0 (J = 0.006955)	1.173 for all fractions	49.0
DSDP 296-63-4	lapilli tuff	51.8	47.9 ± 4.0 (72%)	47.5 ± 4.0 (J = 0.010201)	1.377 for all fractions	47.5

* $\lambda_e = 0.585 \times 10^{-10} \text{ yr}^{-1}$, $\lambda = 5.305 \times 10^{-10}$, $^{40}\text{K}/^{39}\text{K} = 1.19 \times 10^{-4}$

†Total age = $(1/\lambda) \ln \{1 + J(^{40}\text{Ar}-^{39}\text{Ar})_{\text{total}}\}$, where $(^{40}\text{Ar}-^{39}\text{Ar})_{\text{total}} = \sum f_i (^{40}\text{Ar}-^{39}\text{Ar})_i$, f_i = fraction of ³⁹Ar released.

‡Defined for the fraction of ³⁹Ar released, which is indicated in a bracket.

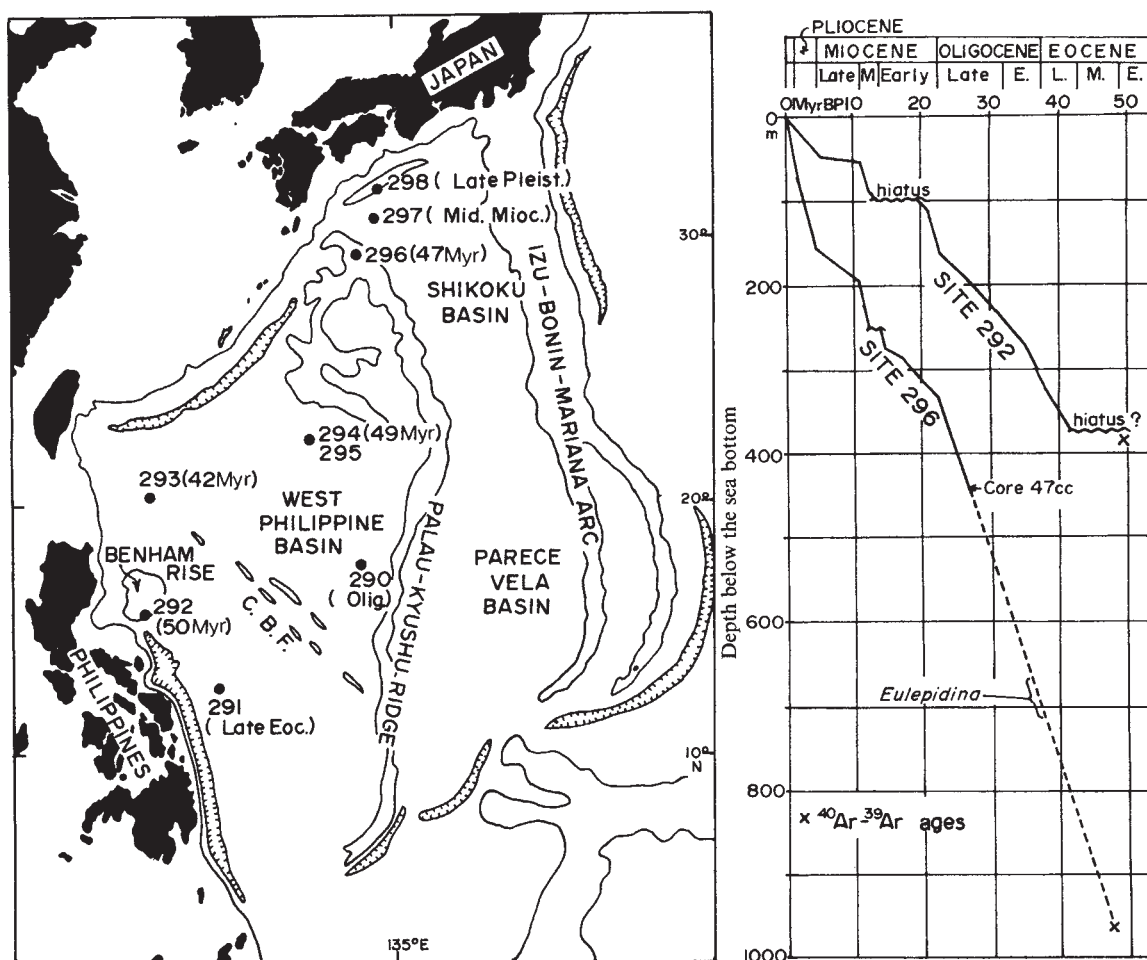
§Determined for the experimental data (temperature fractions) by means of York's least square fitting method.

||The isochron and plateau ages were determined only for three lowest temperature fractions; 600, 700, 800 °C. Hence, the age may not be very reliable.

adopted, the age would be approximately 41 Myr BP. An ⁴⁰Ar-³⁹Ar age obtained for a basalt (about 16 m below the chalk sample) gives 47.5 Myr BP (Fig. 1) although this age difference may be due to a depositional hiatus. A K-Ar age obtained by McKee¹² for Core 45 (38.1 Myr BP) and Core 47 (37.1 Myr BP) may reflect Ar-loss or some alteration effects, since the K-Ar ages are even younger than the microfossil ages in the overlying sediments (Core 41) which is about 40 cm above the basalts. Although nannofossils,

were tentatively assigned the lower Pliocene age for the basalt part of the Hole 293, their rarity means that the data are highly questionable. As the basalt sample used for ⁴⁰Ar-³⁹Ar dating is brecciated gabbro¹, the radiometric age is unlikely to represent the basement age. Hole 294 is the northernmost drilling site in Philippine Sea and also furthest from the CBF. If the CBF was truly a spreading ridge axis, one would expect to find the oldest age at this site, but this is not the case.

Fig 1 a, DSDP Leg 31 sampling sites: ⁴⁰Ar-³⁹Ar ages (Myr) and minimum basement ages estimated by microfossils are shown in a bracket (left). b, the fossil dates and ⁴⁰Ar-³⁹Ar ages are plotted against a depth below the sea bottom for site 292 and 296, which are only sites in Leg 31 to show a reliable basement fossil age.



Summarising, the western Philippine Basin is very likely to be Tertiary in age, there is no positive age evidence to suggest the spreading ocean ridge origin of CBF and the Philippine Basin is significantly younger than the currently proposed magnetic anomaly lineations would imply⁴⁻⁸.

Several authors¹³⁻¹⁵ have identified magnetic anomaly lineations in Shikoku Basin suggesting a spreading origin for the Basin. Using Leg 31 data, Watts and Weissel¹⁶ estimated the oldest age of the basement to be 24 to 27 Myr BP, and that the east-west-directed spreading of Shikoku Basin started about the same time. Only one ⁴⁰Ar-³⁹Ar age has been obtained for lapilli tuff at site 296, located on the Kyushu-Palau Ridge, and this is considerably older than the assumed basement age¹⁸. Site 296 shows gradational change from purely lapilli tuff to pure clastic rock (mainly nannofossil chalk); in Core 38, *Eulepidina*, a larger foraminiferan, indicated shallow water habitat around the middle (Core 56 and 57) of this changing interval¹. These results suggest that the pyroclastics accumulated in a rather shallow sea which then subsided. An average sedimentation rate of about 2.72 cm per 1,000 yr can be estimated for the time interval from 22.5 Myr BP (M/O boundary detected by planktonic foraminifera) to 47.5 Myr BP (⁴⁰Ar-³⁹Ar age) (Fig. 1) which is significantly faster than that of the overlying sediments. The ⁴⁰Ar-³⁹Ar age result is, therefore, in accordance with that expected from the sedimentation rate. The difference between the ⁴⁰Ar-³⁹Ar age and the assumed basement age indicates that either the opening of the Shikoku Basin was a much later event than the formation of the Kyushu-Palau Ridge, or, that the identification of the magnetic anomaly lineation is in error. To resolve the question more radiometric ages of the basement rocks in Shikoku Basin are needed.

The ⁴⁰Ar-³⁹Ar age of the lapilli tuff, seems to favour a suggestion¹³⁻¹⁶ that Shikoku Basin, as well as Parece Vella Basin, was formed by rifting of the Izu-Bonin-Mariana Islands Arc. If the Shikoku Basin was formed this way DSDP site 296, now on the Kyushu-Palau Ridge, was once located on the Izu-Bonin-Mariana Island Arc, or else there should be some similarity between the basement rocks of the Kyushu-Palau Ridge represented by Site 296 materials and of the Izu-Bonin-Mariana Island Arc. In the latter augite-orthopyroxene andesite from Hahajima of the Bonin islands gave a K-Ar age of 40.1 ± 1.2 Myr BP (ref. 17) and dacite from Saipan of the Mariana Islands yielded an age of 45 Myr BP. Both the ages are consistent with larger foraminifera evidence from these islands and similar to that for site 296 (ref. 18). In addition Meijer¹⁹ found Pb isotopic ratios in andesite from the Site 296 which were quite similar to those found in volcanic rocks from Iwo-Jima, south of the Bonin Islands, both were very different from typical oceanic basalt Pb isotopic compositions, but were similar to those from the island arcs. These facts, therefore, indicate genetic similarities between site 296 basement materials and the Izu-Bonin-Mariana island arc, favouring the rifting apart theory of the origin of the Shikoku Basin.

M. OZIMA
I. KANEOKA

Geophysical Institute,
Department of Paleontology,
University of Tokyo,
Tokyo 113, Japan

H. UJIE

National Science Museum,
Department of Paleontology,
Tokyo 160, Japan

Received 3 March; accepted 27 April 1977.

¹ Deep Sea Drilling Project, Init. Report, Leg 31 (1975).

² Karig, D. E. *J. geophys. Res.* **76**, 2542 (1971).

³ Saito, K. & Ozima, M. *Earth planet. Sci. Lett.* **33**, 353 (1977).

⁴ Uyeda, S. & Ben-Avraham, Z. *Nature phys. Sci.* **240**, 176 (1972).

⁵ Segawa, J. *Kagaku* (in Japanese) **43**, 45 (1973).

- ⁶ Karig, D. E., Ingle, J. C., Bouma, A. H., Ellis, H., Haile, N., Koizumi, I., MacGregor, I. D., Moore, J. C., Ujiie, H., Watanabe, T., White, S. M., Yasui, M. & Yi-Ling, H. *Science* **246**, 458 (1973).
- ⁷ Loudon, K. E. *A.G.U. Geophysical Monograph* **19**, 253 (1976).
- ⁸ Watts, A. B., Weissel, J. K. & Larson, R. L. *Tectonophysics* **37**, 167 (1977).
- ⁹ Blow, W. H. *Proc. 1st Int. Conf. Plankt. Microfossils Geneva* **1**, 197 (1969).
- ¹⁰ Berggren, W. A. *The Micropaleontology of Oceans* **693** (Cambridge University Press, Cambridge, 1971).
- ¹¹ Berggren, W. A. & Couvring, J. Van *Paleogeogr. paleoclimat. Paleocol.* **16**, 1 (1974).
- ¹² McKee, E. H. *Init. Rep. DSDP* **31**, 599 (1975).
- ¹³ Murauchi, S., Asanuma, T. & Saki, K. *Kaiyo-Kagaku* (in Japanese) **6**, 56 (1974).
- ¹⁴ Watts, A. B., Weissel, J. K. & Larson, R. L. *Tectonophysics* **37**, 167 (1977).
- ¹⁵ Tomoda, Y., Kobayashi, K., Segawa, J., Nomura, M., Kimura, K. & Saki, T. *J. geomag. Geoelectr.* **28**, 47 (1975).
- ¹⁶ Watts, A. B. & Weissel, J. K. *Earth planet. Sci. Lett.* **25**, 239 (1975).
- ¹⁷ Kaneoka, I., Isshiki, N. & Zashu, S. *Geochim. J.* **4**, 53 (1970).
- ¹⁸ Ujiie, H. & Matsumaru, H. *Memoir of National Science Museum* (in the press).
- ¹⁹ Meijer, A. *Geol. Soc. Am. Bull.* **87**, 1358 (1976).
- ²⁰ York, D. *Earth planet. Sci. Lett.* **5**, 359 (1969).

Alpha-induced coloration reversal in biotite

USING previously developed techniques for the investigation of radiation coloration in biotite, we have discovered the hitherto unknown fact that the onset of reversal occurs precisely at the end of the α range and is a phenomenon independent of the darkening peak.

Since the earliest original investigations of Joly¹ into the nature of pleochroic haloes, it has been known that α radiation can produce a 'reversal' effect in biotite (sometimes called 'inversion'). The usual effect of natural or artificial α irradiation is to produce a characteristic darkening of the biotite. This darkening increases with dose until a certain point (that is, the optical absorption coefficient saturates) and then the mineral begins to lighten as the dose is further increased. With sufficient dosage it is possible to lighten the biotite to the point where it actually becomes less dark than its naturally occurring shade.

The most quantitative study of reversal was carried out by Deutsch^{2,3} and his collaborators. Their technique involved the irradiation of small flakes of biotite with non-collimated radon α particles. They found that in their experimental configuration reversal took place at doses over approximately 10^{17} alphas cm^{-2} which is about four orders of magnitude beyond the onset of darkening. They also determined that the decrease in darkening is a much slower phenomenon than the increase in darkening near saturation. Other important halo investigators before and after Deutsch have also remarked on the reversal phenomenon^{4,5}.

Using techniques of Van de Graaff irradiation of biotite and precision photomicrography which have been described elsewhere⁶, we have discovered a hitherto unknown feature of reversal. In the literature it is implicit that reversal occurs at the point of maximum darkening or, at least, is related to the ionisation which causes the darkening peak. We have found that in fact reversal does not begin at the darkening peak, but at the foot of the peak at a point corresponding to the α -extrapolated ionisation range in biotite which is $35.1 \mu\text{m}$ at 7.69 MeV (ref. 6).

The most significant conclusion from these results is that the reversal process is associated with the depositing of α -particles at the end of their range and is not an effect due to ionisation, as is the darkening. In other words, the physical mechanisms of darkening and reversal seem to be quite different.

Figure 1 shows the optical absorption coefficient for four successive α -irradiations of the same piece of biotite. (As in ref. 6 we used a specimen from the Faraday pegmatite in Bancroft, Ontario, Canada.) For clarity we used the maximum ²³⁹U α -family energy, 7.69 MeV. The successive development of the reversal is readily apparent. The minimum of reversal corresponds within the error of measurement to the extrapolated ionisation range as given above. Our work, based on the results shown in Fig. 1 and other bombardments, indicates a reversal onset at approximately 5.2×10^{14} alphas cm^{-2} . We have found a plateau area of at

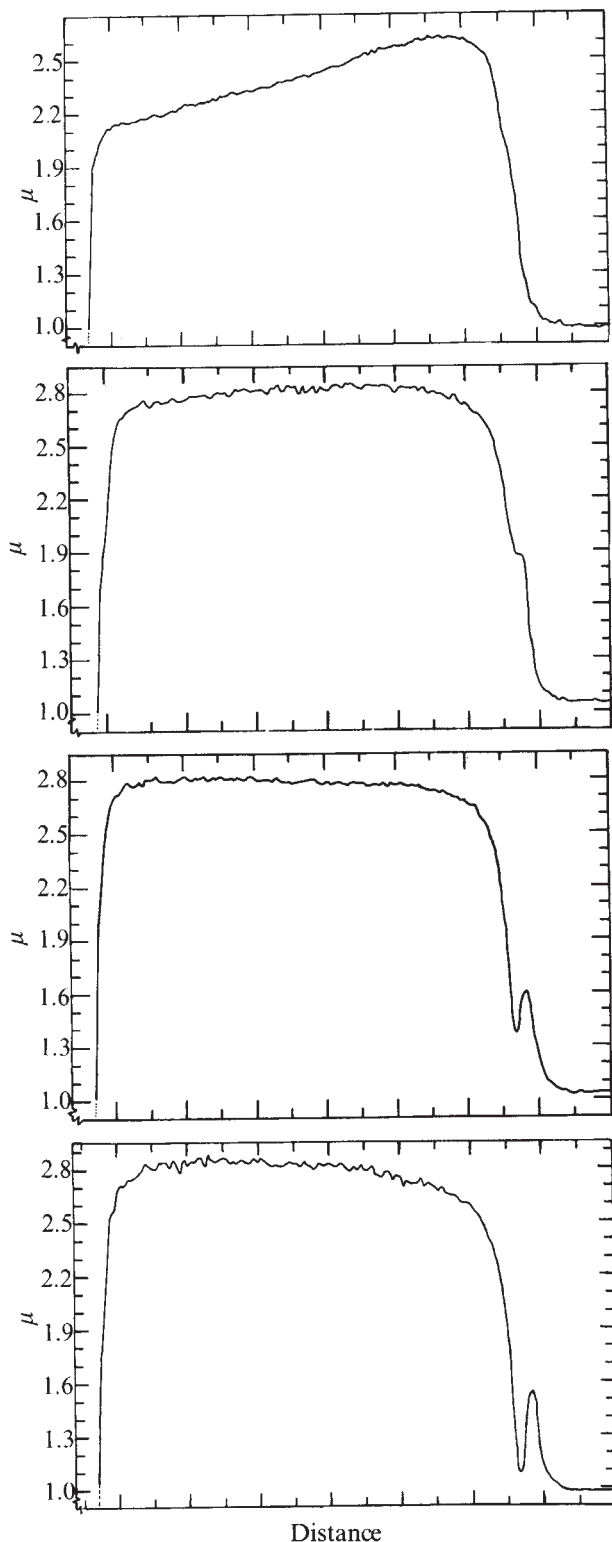


Fig. 1 These curves are densitometer tracings, suitably processed according to the method of ref. 6, of single-energy α bombardments of biotite. For each curve the ordinate is the optical absorption coefficient (in units of the natural coefficient for the biotite so that $\mu = 1$ is the unirradiated mineral). All four energies were 7.69 MeV and the doses, from top to bottom were approximately 2.6×10^{14} alphas cm^{-2} , 10^{15} alphas cm^{-2} , 2.1×10^{15} alphas cm^{-2} , and 4.0×10^{15} alphas cm^{-2} . Note the distinct separation between the maximum coloration at $32.2 \pm 0.2 \mu\text{m}$, and the reversal at the extrapolated ionisation range of $35.1 \pm 0.2 \mu\text{m}$.

most one order of magnitude between saturation and the onset of reversal. Our onset of reversal dose is several orders of magnitude below that found by Deutsch and our plateau area somewhat smaller. We do not consider that these conflict with Deutsch's results considering the precision of our work and his semi-quantitative approach. We repeat, however, that the reversal onset does not occur at the same place along the α track as does the darkening peak. The latter occurs at the point of maximum ionisation⁶.

Previous explanations of reversal go back to Joly who believed the halo development process to be a photographic-like one with a latent image being 'developed' by the α radiation. He attributed the reversal to some unspecified 'photographic' effect. Henderson and Turnbull⁴ made Joly's conjectures more specific by attributing the reversal to an Eberhard-like effect in the biotite. In the development of a negative the Eberhard effect causes regions adjacent to sharply defined dense areas to become lighter than they should due to a diffusion into the affected regions of an excess of soluble bromide. This misinterpretation of the Eberhard effect caused Henderson and Turnbull to underestimate the alpha ranges when determined from reversal haloes⁴. As our Fig. 1 shows, the reversal effect has a distinct and sharply defined onset and is not at all characteristic of an effect diffusing from the darkening peak outward.

Deutsch and Janssens also considered the geological phenomenon of metamictisation and concluded after some simple calculations that this could not be the cause of reversal.

The phenomenon of reversal does not seem well known in solid state physics. It would be of interest to α -bombard some well understood system, such as an alkali halide, to see if colour reversal could be obtained.

We thank R. Ward for providing us with biotite specimens and A. Barrett who provided access to and assisted with the densitometer. We also thank A. M. de Graaf for helpful discussions.

CYRUS MOAZED

Harry Diamond Laboratories,
2800 Powder Mill Road,
Adelphi, Maryland 20783

ROBERT OVERBEY
RICHARD M. SPECTOR

Physics Department,
Wayne State University,
Detroit, Michigan 48202

Received 12 April; accepted 22 April 1977.

¹ Joly, J. *Phil. Trans. R. Soc. A* **217**, 51-79 (1917).

² Deutsch, S., Kipfer, P. & Picciotto, E. *Nuovo Cim.* **VI**, 796-810 (1957).

³ Deutsch, S. & Janssens, P. *Nuovo Cim.* **XI**, 473-483 (1959).

⁴ Henderson, G. H. & Turnbull, L. G. *Proc. R. Soc. A* **145**, 582-591 (1934).

⁵ Gentry, R. V. *Ann. Rev. Nuc. Sci.* **23**, 347-462 (1973).

⁶ Moazed, C., Overbey, R. & Spector, R. M. *Nature* **258**, 315-317 (1975).

The Thera eruption and Late Minoan-IB destructions on Crete

MARINATOS' theory¹ of the decline of the Minoan civilisation by paroxysmal volcanic events of the Thera volcano is widely accepted, and many papers on the topic have been published, especially since the International Scientific Congress on the Volcano of Thera². The main and controversial problem is the fact that the simultaneous devastations of numerous Late Minoan (LM) IB sites and palaces in Crete are dated archaeologically (concerning the development of Minoan pottery) as 30-50 yr later than the destruction of the important LM settlement on Santorini (Acrotiri ruins)³. Following Marinatos' theory, that the devastations in Crete were caused by the Thera volcano, two eruptions (if

chronology is correct) must have occurred. The first, destroying the LM-IA settlements on Santorini, in about 1500 bc, the second, causing the LM-IB destructions in Crete, about 1450 bc. According to a recent theory, only LM-IA destructions on Crete are connected with the Thera outburst^{4,5}. We report here fieldwork on Santorini which shows clearly that neither products of two volcanic outbursts nor of a 'two-phase' eruption⁶, differing in a time interval of about 50 yr, exist, and that the collapse of the Santorini caldera, which caused devastating earthquakes and tsunamis in Crete, did not occur about 50 yr after the volcanic paroxysm.

If two eruptions had occurred, volcanic products of both would exist. In that case, the time interval between both volcanic events should be manifested by an erosional disconformity and a weathered layer dividing the products of the two eruptions. From the volcanological point of view, the experts are unanimous that the Upper Pumice Series⁷, which covers the destroyed LM settlements on Santorini, was deposited during a single volcanic phase. We carried out new fieldwork on Santorini in autumn 1976 which confirmed this current view. According to the stratigraphy and lithologic character of the Upper Pumice Series⁸ the LM outburst began with catastrophic earthquakes that, after an interval between some weeks and 2 yr (ref. 2), were followed by a large pumice-fall. This pumice-fall layer is overlain by base surge deposits⁹, which locally show a slight erosional break. It should be stressed, however, that this break represents only a very short span of time of volcanic inactivity, that is of some days up to at most few weeks. The eruption culminated in some violent ash-flow eruptions by which the bulk of the Upper Pumice Series was deposited. These eruptions were connected with the beginning of the collapse of the central part of the old

Santorini island. This is evidenced by abundant huge lithic blocks, as much as 3 m in diameter, embedded in the ash-flow material. These blocks were formed by a falling-in of parts of the old volcanic pile along circular and/or polygonal fractures. Still in the course of the collapse the blocks were ejected immediately by simultaneous paroxysmal eruptions. This evidence is inconsistent with Doumas' theory¹⁰ of a formation of the Santorini caldera about 50 yr after the Thera outburst, that is around 1450 bc. This theory should explain the time interval between the devastations both on Santorini and in Crete.

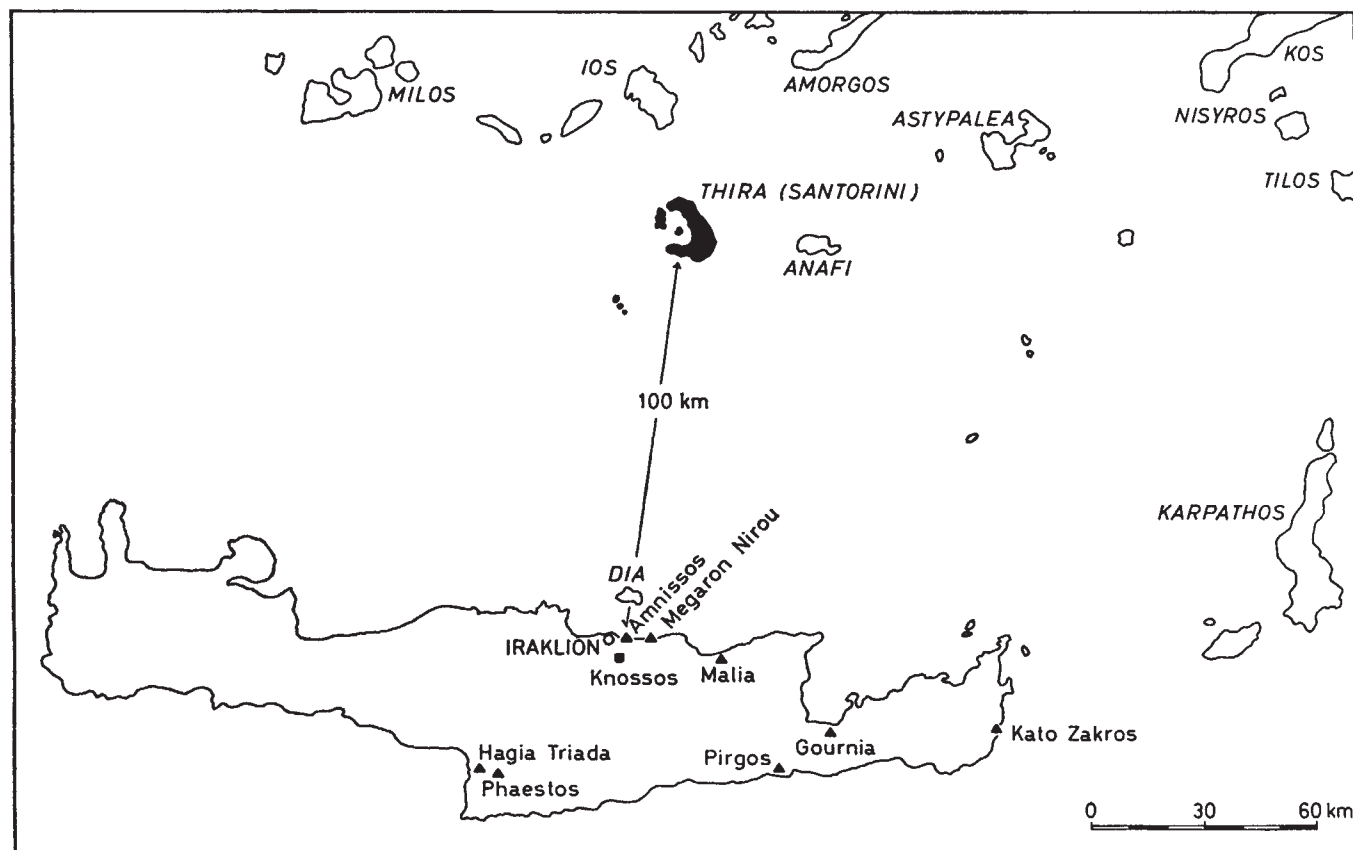
According to our investigations on Santorini and Crete, the devastations in Crete around 1450 bc can be explained only by the following alternative. (1) The archaeological dates must be revised. The destructions on Santorini and in Crete occurred simultaneously (both around 1500 bc or both around 1450 bc).

From the archaeological point of view, a modification of the ceramic chronology is not supported and would, moreover, involve insuperable difficulties. (2) The devastations in Crete were caused neither by the eruptions of the Thera volcano nor as a result of the subsidence of the Santorini caldera.

The results of our fieldwork are inconsistent with the theory that the LM-IB devastations in Crete were the consequence of a volcanic disaster on Thera. According to this theory^{11,12,13,14} the LM-IB devastations on Crete have been caused by tephra-fall and/or tsunamis and/or catastrophic earthquakes, connected with the Thera outburst or the Santorini caldera collapse.

Tephra-fall: Our fieldwork showed that visible remnants of the LM tephra-layer in central and eastern Crete do not exist. Special attention was paid to localities in which such remnants should have been preserved, for example in

Fig. 1 The most important Late Minoan sites on Crete and the location of Thira (Santorini) and other Cycladian islands.



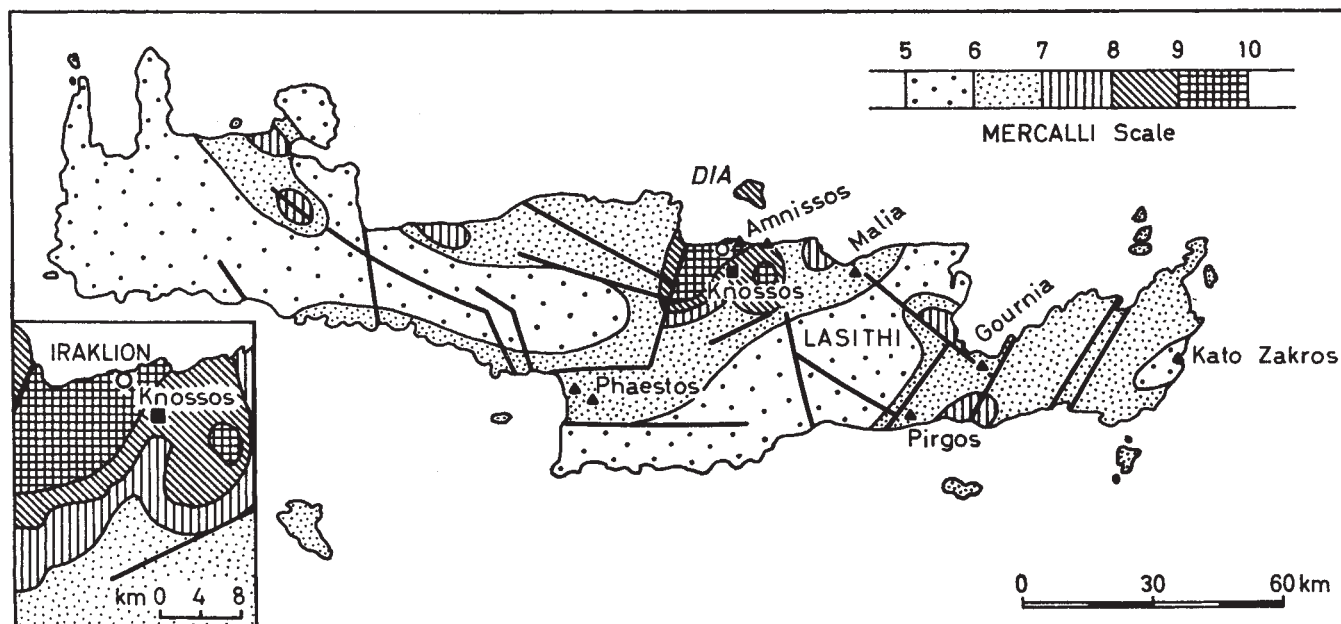


Fig. 2 Isoseismic sketch map of Crete of the great regional earthquake from June, 26, 1926. The relation of the earthquake intensity to the block structure of Crete is clearly visible. (According to Sieberg²¹.)

recent soil profiles at the northern coast, exposed by deep open cuts and drain channels or in crevices and sinks in the limestone mountains around the Lasithi plateau. Residues of some washed soil samples (from Pargos and Malia) contain few small grains of volcanic origin, that is, feldspar, volcanic glass ($n \sim 1.510$, which corresponds to the Upper Pumice Series on Thera⁸), and rare pyroxene. Their meagre quantity, however, is inconsistent with a suggested thickness of the tephra-layer of up to 10 cm (ref. 6) or even of 1 cm (ref. 15). Quantities of 50–750 volcanic particles per cm^3 as indicated by Boeschoten¹⁵, have never been found. The discrepancy between the refractive index of the glass shards in Boeschoten's samples, that is, $n \sim 1.503$, and the index of the volcanic glass of the Upper Pumice Series ($n \sim 1.510$, ref. 8), was already stated¹⁶. In other publications dealing with the presence of LM volcanic tephra on Crete^{16,17} is reported that only traces of such tephra have been found, the highest concentrations only in Minoan settlements.

There is a simple explanation for such a presence of pumice particles in these settlements, namely the use of pumice in workshops and in households. Lumps of pumice were already used by Minoan people for various purposes. This is illustrated by findings of numerous lumps and blocks of pumice (up to 30 cm in diameter) in Malia, Gournia, Knossos and Kato Zakros. Because of its size this pumice could not have reached Crete by way of an even paroxysmal volcanic eruption on Santorini, Milos or Nisyros. This material was mainly imported from these islands. Pieces from Malia, Knossos and Kato Zakros are characterised by grooves and traces of working.

The volcanic ash horizons in eastern Mediterranean cores, which were attributed to the Thera outburst and considered as an evidence of an important ash-fall in Crete and the surrounding Aegean Sea¹¹, are of three or four different ages¹⁸. The abnormal local thickness (as much as 212 cm) (ref. 11) of ash-layers in some cores is due to post-depositional mixing and transport. Furthermore, new problems concerning the stratigraphic correlation of these ash beds in the eastern Mediterranean cores have been pointed out¹⁹. The distribution chart of the tephra from the LM

Thera outburst, "imaginative but very uncritically"¹⁸ constructed by Ninkovich and Heezen¹¹, must therefore be discarded as one of the chief arguments for a disastrous tephra-fall on Crete.

Tsunamis: According to our observations in Amnisos, Megaron Nirou, Malia, Gournia and Kato Zakros the force of tsunamis originated by volcanic or seismic events in connection with the Thera outburst must have been more shallow than previously supposed¹⁴. According to our calculations, the uppermost height of the tsunamis amounts to only 8–10 m (ref. 20).

Volcanic earthquakes: The devastations in Crete about 1450 BC could not have been caused by volcanic earthquakes because the Thera outburst occurred about 50 yr earlier.

We conclude from the above that all non-human devastations around 1450 BC in Crete, that is, that of Amnisos, Megaron Nirou, Malia, Gournia, Kato Zakros on the northern and eastern coast and that of Pargos, Phaistos and Hagia Triada on or near the southern coast of Crete (Fig. 1)²⁰, have been caused by a catastrophic regional earthquake, similar to that of AD 1926 (refs 14, 21). The complicated block structure of Crete (Fig. 2) is the reason that, as during the great recent earthquakes (1926, 1856, 1846, 1815, 1780 and so on)²¹, tectonically unstable regions have violently been affected. More stable regions were only slightly touched. This regionally different behaviour of Crete with respect to earthquakes is consistent with the fact that the palace of Knossos was not demolished during the regional earthquake around 1450 BC. The same phenomenon was observed in 1926, when Evan's reconstructions of Knossos did not suffer but the neighbouring Iraklion was strongly demolished.

According to our observations there is no connection between the decline of the Minoan civilisation and the Thera outburst. This volcanic paroxysm occurred around 1500 BC and did not have a revival around 1450 BC, at a time of the great devastations in Crete. All speculations that the Thera outburst should have caused 'a large natural catastrophe in three stages', that is the plagues of Egypt, the deluge of Deukalion and the submergence of Atlantis¹⁴

are imaginative but, from the scientific point of view, not convincing.

HANS PICHLER

Mineralogisch-Petrographisches
Institut der Universität
Wilhelmstrasse 56
D-7400 Tübingen

WOLFGANG SCHIERING

Seminar für Klassische Archäologie der Universität
Schloss
D-6800 Mannheim 1, FRG

Received 11 March; accepted 11 May 1977.

- ¹ Marinatos, S. *Antiquity* 13, 425–439 (1939).
- ² Marinatos, S., ed. *Acta Internat. Sci. Congr. on the Volcano of Thera* (Athens, 1971).
- ³ Marinatos, S. *Excavations at Thera I-IV* (Athens, 1968–1973).
- ⁴ Hood, S. *Acta Internat. Sci. Congr. on the Volcano of Thera*, 377–383 (Athens, 1971); *Acta 3rd Cretological Congr.* 1, 111–118 (1976).
- ⁵ Wilson, A. L. *Am. J. Archaeol.* 80, 419–432 (1976).
- ⁶ Luce, J. V. *Am. J. Archaeol.* 80, 9–18 (1976).
- ⁷ Pichler, H. & Kussmaul, S. N. *Jb. Mineral Abh.* 116, 268–307 (1972).
- ⁸ Günther, D. & Pichler, H. N. *Jb. Geol. Paläont. Mh.* 394–415 (1973).
- ⁹ Pichler, H. *Die Naturwiss.* 60, 198 (1973).
- ¹⁰ Doumas, C. *Antiquity* 48, 110–115 (1974).
- ¹¹ Ninkovich, D. & Heezen, B. C. *Colston Papers*, 17, 413–452 (Butterworths, London 1965).
- ¹² Page, D. L. *Soc. Promotion Hellenic Studies*, Suppl. Pap. No. 12 (London, 1970).
- ¹³ Warren, P. *Gnomon* 45, 173–178 (1973).
- ¹⁴ Galanopoulos, A. G. *Acta Internat. Sci. Congr. on the volcano of Thera*, 184–210 (Athens, 1971).
- ¹⁵ Boeckschoten, G. J. *Opera Botanica* No. 30 (Evolution in the Aegean), 40–48 (Lund, 1971).
- ¹⁶ Vitaliano, C. J. & Vitaliano, D. B. *Am. J. Archaeol.* 78, 19–24 (1974).
- ¹⁷ Cadogan, G., Harrison, R. K. & Strong, G. E. *Antiquity* 46, 310–313 (1972).
- ¹⁸ Olausson, E. *Opera Botanica* No. 30 (Evolution in the Aegean), 29–39 (Lund, 1971).
- ¹⁹ Pichler, H. & Friedrich, W. *Nature* 262, 373–374 (1976).
- ²⁰ Pichler, H. & Schiering, W. (in preparation).
- ²¹ Sieberg, A. *Denkschr. mediz.-naturw. Ges. Jena*, 18/2 (Jena 1932).

Charge stabilisation of very small particles of mixed oxides

IN our laboratory, a rock salt type structure has been found for mixed oxides of the type $Me(II)_xCr(III)_{2/3(1-x)}O$. $Me(II)$ can be Co, Ni, Cu, and Zn or mixtures of these. Other authors^{1,2} have found similar compounds of the general formula $Me(II)_xMe(III)_{2/3(1-x)}O$. We suggest a structure where a very small particle, of the order of 2–10 nm, has been stabilised by a surplus of O^{2-} ions on the surface, compensating an excess positive charge uniformly distributed in the bulk.

Besides the presence of $Me(II)_xCr(III)_{2/3(1-x)}O$, X-ray analysis shows MeO in varying small amounts. The particle size, as determined by X-ray diffraction, is typically in the order of 3–10 nm. In addition, there is a size dependence of the lattice parameter such that a small particle will have a relatively smaller lattice parameter of 412 pm, whereas for a 10 nm particle it will be around 420 pm.

From the structural data, one may conclude that the cations are randomly distributed on the octahedral sites in the crystal. Any ordered arrangements of these, or occupation of other sites, would destroy the rock salt structure. From electron diffraction and microscopy, it seems that in particular the (111) face is towards the vacuum side. In the zinc case, epitaxial ZnO (001) on $Zn_xCr_{2/3(1-x)}O$ has been observed (Fig. 1). Below 400 °C, the structure is stable in wet and dry atmospheres. Above this temperature, spinel phase and/or MeO/Cr_2O_3 is formed in an irreversible process.

As a possible explanation of these observations, we suggest a model where the particle surface has become negatively charged. If we, for simplicity, consider a finite linear chain of Me^{2+} , O^{2-} ions, the energy per O^{2-} ion would be lower than in the corresponding infinite chain. Substitution of 2 Cr^{3+} for 3 Me^{2+} and rearrangement of the chain so that both endpoints are occupied by O^{2-} ions results in a more stable configuration where the energy per O^{2-} ion is higher than in the infinite. The energy depends inversely on the chain length. For a bulk crystal with rock salt structure we may perform a similar process. Let us consider a small cube of NiO with an edge

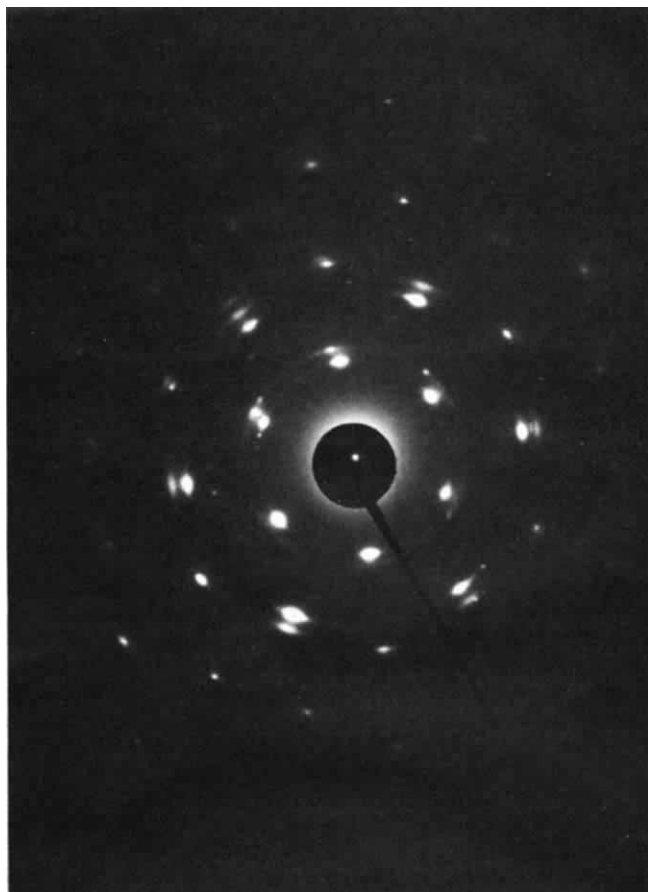


Fig. 1 Transmission electron diffraction patterns of a mixture of ZnO and $Zn_xCr_{2/3(1-x)}O$. The intense 6-fold symmetry spots arise from the (001) plane in ZnO while the less intense 6-fold symmetry spots are due to the (111) planes of the mixed oxide phase.

length of $\sim na$, where a is the lattice parameter ($= 418$ pm). The cube contains $4n^3$ NiO pairs; $12n^2$ of the sears located on the surface, edges and corners of the cube. Substitution of 2 Cr^{3+} in this lattice for 3 Ni^{2+} creates one cation vacancy. The surface acts as a sink for the vacancy and eventually 1 Ni^{2+} will be removed from the surface resulting in a negative charge of $-2e$ on the surface. It should be added that the Madelung energy of the Ni^{2+} is higher in the bulk than on the surface, making this process favourable. Assuming that the process is favourable as long as there are cations left on the surface, we end up with a crystal with

$$\begin{aligned} 4n^3 O^{2-} \\ 24n^2 Cr^{3+} \\ 4n^3 - 32n^2 Ni^{2+} \end{aligned}$$

so

$$x = 1 - \frac{9}{n}$$

For the size mentioned (~ 5 nm), n is about 12 corresponding to $x \approx 0.25$. Thus for small n the Cr^{3+} concentration becomes appreciable. On the other hand, this will result in a smaller lattice parameter, since the ion radius of Cr^{3+} is less than the ion radius of Me^{2+} . Two other energy contributions may stabilise this arrangement. Firstly, the smaller lattice parameter will result in a larger Madelung energy. Secondly, the Cr^{3+} ion has a sizeable octahedral configuration energy.

In the case of an octahedron, the surface-to-bulk ratio is higher than for the cube. Besides, we may construct the crystal so that it has the negatively charged (111) planes towards the vacuum side. Electron diffraction indicates that this is a likely arrangement (Fig. 1).

In order to estimate the energy changes in such an arrangement one should calculate, in addition to the Madelung energy, the attractive energy from the bulk surface interaction, as well as the repulsive energies from the negatively charged surface and from the positively charged bulk. This is not easily done, for a cube or for an octahedron. An estimate of the relative magnitudes involved, however, can be obtained by considering a 'spherical crystal' where the excess positive charge in the bulk is compensated for by a corresponding negative charge on the surface. We are not, in this argument, considering the repulsive potential between nearest neighbour ions, since this should be relatively independent of the various configurations we are considering.

Let N be the number of O^{2-} ions. The maximum number of O^{2-} in the surface is then given by $kN^{2/3}$.

We then have $2k N^{2/3} Cr^{3+}$ ions and $N - 3kN^{2/3} Ni^{2+}$ ions. The surface has a total negative charge of $-2kN^{2/3} e = -Q$. Similarly, the bulk has an excess positive charge of $+Q$.

Using the superposition principle, one may calculate the electrostatic energies involved. Firstly, there is a Madelung energy from the lattice, being approximately

$$U_M = -\frac{\alpha}{a} \frac{1}{4\pi\epsilon_0} 4e^2 \left\{ \left(N - 3kN^{2/3} \right) + \frac{3}{2} 2kN^{2/3} \right\} \\ = -\frac{4\alpha}{a} \frac{e^2}{4\pi\epsilon_0} N \quad (1)$$

where $\alpha \approx 3.50$ is the Madelung constant and the first term represents MeO , the second (approximately) Cr_2O_3 .

Furthermore, there are the Coulomb energies from this particular charge arrangement.

$$U_C \approx \frac{Q^2}{4\pi\epsilon_0 r} \left(\frac{1}{2} + \frac{3}{5\epsilon} - \frac{1}{\epsilon} \right) = \frac{Q}{4\pi\epsilon_0 r} \left(\frac{1}{2} - \frac{2}{5\epsilon} \right) \quad (2)$$

the three terms in the bracket representing the surface repulsive energy, the bulk repulsive energy, and the attractive energy from the surface bulk interaction; ϵ is the dielectric constant. Inserting the value of Q and using $r = k_1 N^{1/3} a$, we find

$$U_C = \frac{k^2 N^{4/3} e^2}{4\pi\epsilon_0 r} \left(\frac{1}{2} - \frac{2}{5\epsilon} \right) \approx \frac{2.6 k^2 N}{4\pi\epsilon_0 a} e^2 \left(\frac{1}{2} - \frac{2}{5\epsilon} \right) \quad (3)$$

The maximum value of k is of the order three, whereas values of ϵ for very small particles are not known, but are most likely to be greater than one. The Coulomb term as such is then positive. In the limit for small spheres $\epsilon \rightarrow 1$ and

$$U_C \approx \frac{k^2}{4} \frac{Ne^2}{4\pi\epsilon_0 a} \quad (4)$$

The stability requirement is that the pair energy should be lower than for a pair in a cube of a similar size. For pure MeO , this is the stable configuration. The most stable configuration is obtained by minimising the expression for the total energy per O^{2-} ion with respect to N and k . Madelung energies of small cubes have been calculated by Christov³. Here we will use the approximative expression

$$U_0 \approx -\frac{4\alpha}{a_0} \frac{e^2}{4\pi\epsilon_0} \left\{ N - 0.24 N^{2/3} \right\} \quad (5)$$

where a_0 is the lattice parameter of pure MeO , and the last term in the bracket represents the smaller energy of surface and edge pairs.

A strongly charged surface will bind polar molecules such as water very strongly. A bonding energy of 200 kJ mol^{-1} will contribute about $-0.75 k N^{2/3} e^2 / 4\pi\epsilon_0 a$

Including this, we obtain for the energy in the case of $k = 2$

$$U \approx \frac{-e^2}{4\pi\epsilon_0 a} N(14 - 1.0 + 1.5 N^{-1/3}) \quad (6)$$

The stability region is determined by $U/U_0 \geq 1$.

$$\text{Thus } \frac{a_0}{a} \frac{13.0 + 1.5 N^{-1/3}}{14 - 3.4 N^{-1/3}} \geq 1 \quad (7)$$

The ratio a_0/a is a function of N . It represents the gain in Madelung energy by contraction of the lattice. Using Vegard's law and suitable values for ion radii we obtain in the present approximation $N \leq 1000$.

For $k = 1$, we obtain $N \leq 30,000$. If, instead, we used the pair energy of the infinite crystal as a stability limit, we would obtain $N \leq 600$ for $k = 1$. Thus, the model qualitatively explains the relative stability of small particles of mixed oxides. A similar situation could exist for other systems. It might also explain the growth of octahedral crystals of $NaCl$ and KCl under various circumstances.

I thank Jørgen Villadsen, who introduced me to the structure and properties of the mixed oxides, and to Haldor Topsøe A/S for permission to publish this work.

P. E. HØJLUND NIELSEN

The Haldor Topsøe Research Laboratories,
Nymollevvej 55,
DK-2800, Lyngby—Denmark

Received 5 April; accepted 27 April 1977.

¹ Laqua, W., Dudda, S. & Reuter, B. *2 anorg. allg. Chem.* **428**, 151–165 (1977).

² Leygraf, C. & Hultquist, G. *Surface Sci.* **61**, 69–84 (1976).

³ Christov, C. *Ann. Physik (Leipzig)* **6**, Folge **15**, 112–123 (1954).

State of iron in glass-like carbon prepared by heat treatment of ferrocene resin

WE have investigated the conversion of organometallic polymers to inorganic materials by heat-treatment. Several of those inorganic materials are industrially useful. Earlier studies have been reported in this journal^{1–3}. We describe here the state of iron dispersed in a glass-like carbon, which was prepared by heat-treatment of an organometallic resin containing ferrocene skeletons as the main chain. This state is unusual and strange both physically and chemically.

The organometallic resin was synthesised by the polycondensation of acetylferrocene and furfural using sulphuric acid as the catalyst⁴. When the molar ratio of furfural (F) and acetylferrocene (A) exceeded $F/A=15$, the resin was not formed. Acetone was added to the starting material which contained furfural at an F/A ratio above 15 (ref. 7). The resins obtained were named 2-resin, 5-resin, 10-resin, 20-resin, 30-resin, 40-resin and 100-resin. The number indicates the number of moles of furfural for 1 mole of acetylferrocene in the starting material. To examine the physical and chemical state of the iron by Mössbauer spectroscopy, 30- and 100-resins were prepared using ferrocene consisting of ⁵⁷Fe enriched to 62.5% (ref. 8). The resin was heated in a vacuum of 1×10^{-3} mmHg from room temperature up to 350 °C for 4 h, then up to 400 °C for 2 h. The heat-treated resin, still in its evacuated container, was

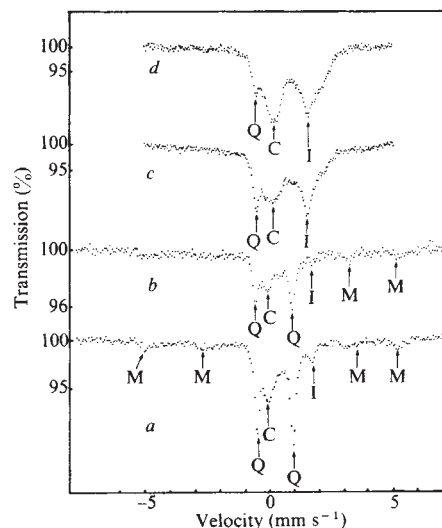
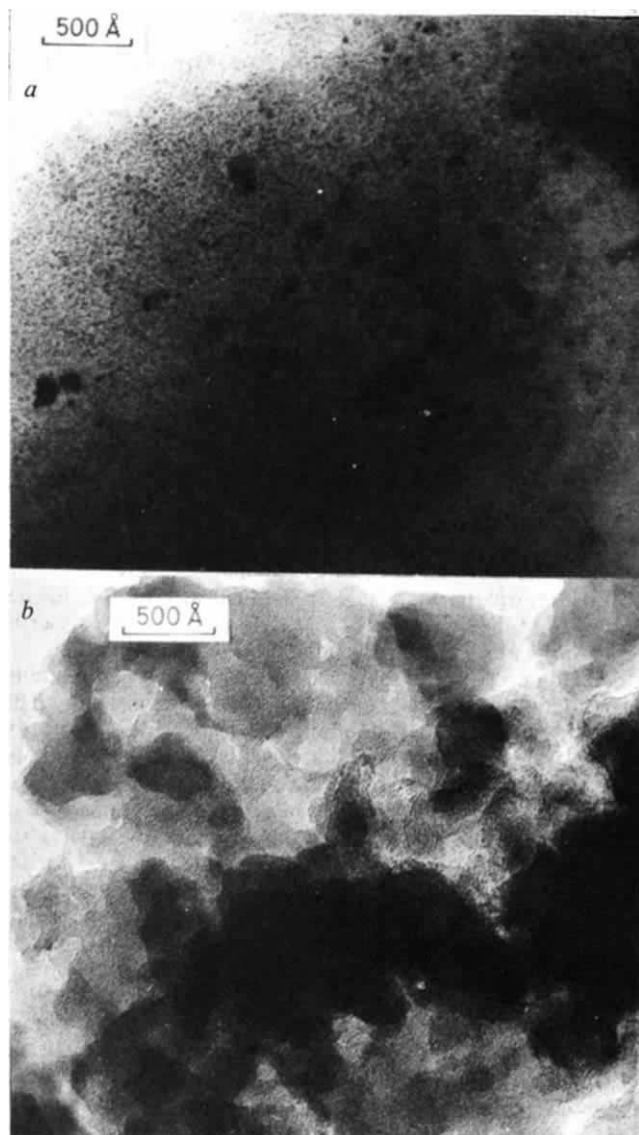
Table 1 Elemental analyses of resins heat-treated at 400 °C

Sample	Fe (wt%)	C (wt%)	H (wt%)	O (wt%)*
2-resin	13.20	69.07	3.68	14.05
5-resin	10.60	71.08	3.22	15.10
10-resin	7.96	72.01	3.58	16.45
20-resin	3.65	73.33	3.65	19.37
30-resin	2.54	75.06	3.77	18.65
40-resin	2.06	76.36	3.69	17.89
100-resin	1.00	77.14	3.78	18.08

* The percentage of oxygen is obtained by subtracting the percentages due to iron, carbon and hydrogen from 100.

placed in a glove box filled with argon; this was used for each measurement.

The resins heat-treated at 400 °C were susceptible to oxidation. In particular, 10-resin was pyrophoric on exposure to air. In the resin heat-treated at 400 °C, therefore, the iron liberated from the ferrocene skeleton would be present as ultra-fine particles or in a low valence state (i.e. Fe^+ or Fe^{2+}). As ferrocene is decomposed by heat-treatment above 400 °C, iron carbide is not formed but metallic iron is produced⁹. So there is probably no iron carbide in the resins heat-treated at 400 °C.

Fig. 1 Electron micrographs of the resins heat-treated at 400 °C, *a*, 2-resin; *b*, 100-resin.**Fig. 2** Mössbauer spectra of the resins heat-treated at 400 °C (4.2 K); *a*, 2-resin; *b*, 10-resin; *c*, 30-resin; *d*, 100-resin.

Infrared spectra showed that the resin structure was decomposed by heat-treatment at 400 °C (ref. 7). The X-ray diffraction patterns indicated that the matrix of heat-treated resin was a glass-like carbon. Results of elemental analysis of the heat-treated resin are shown in Table 1. Electron micrographs of 2- and 100-resins heat-treated at 400 °C are shown in Fig. 1. As shown in Fig. 1*a*, in 2-resin iron particles 10–25 Å in diameter were dispersed uniformly in the glass-like carbon, and larger iron particles (30–150 Å) were thinly scattered. This dispersion was also observed in 5- and 10-resins heat-treated at 400 °C. In 20- and 30-resins heat-treated at 400 °C, iron particles of irregular size (10–150 Å) were dispersed in the glass-like carbon matrix. However, in 40- and 100-resins heat-treated at 400 °C no iron particles were visible, as shown in Fig. 1*b*; the iron particles were evidently below 10 Å in size. In Table 2 are shown the X-ray diffraction data of 2-resin heat-treated at 400 °C (Table 2). The ultra-fine particles included iron of hexagonal close-packed (h.c.p.) ($a_0 = 3.30$, $c_0 = 5.21$) and distorted body-centred cubic (b.c.c.) ($a_0 = 2.95$) type structures. Iron of h.c.p. type structure has previously only been recognised at high pressures of 130 kbar and above¹⁰. These iron particles of h.c.p. and distorted b.c.c. type structures corresponded to those of 10–25 Å in Fig. 1*a*. Diffraction peaks due to the iron particles of 30–150 Å were not observable, since they were small in number. The X-ray diffraction patterns of 5- and 10-resins heat-treated at 400 °C were similar to that of 2-resin. No diffraction peaks due to iron particles were detectable in the other resins heat-treated at 400 °C. Mössbauer spectra of 2-, 10-, 30-, and 100-resins heat-treated at 400 °C are shown in Fig. 2. In the Mössbauer spectrum of 2-resin, four peaks marked M were due to magnetic splitting of ferromagnetic iron. The isomer shift of the two peaks marked Q was $+0.21 \pm 0.03 \text{ mm s}^{-1}$;

Table 2 X-ray diffraction data of ultra-fine iron particles in 2-resin heat treated at 400 °C.

Hexagonal close-packed iron		Distorted body-centred cubic iron	
d (Å)	(hkl)	d (Å)	(hkl)
2.85	(100)	2.09	(110)
2.61	(002)	1.48	(200)
2.51	(101)		
1.61	(110)		
1.48	(103)		
1.39	(112)		
1.30	(004)		
1.10	(203)		

the separation was $1.44 \pm 0.03 \text{ mm s}^{-1}$. These peaks were possibly due to the quadrupole splitting of h.c.p. iron. The peak marked I at $+1.70 \pm 0.03 \text{ mm s}^{-1}$ was attributed to Fe^{+} , because the large isomer shift value agreed with that published for Fe^{+} (ref. 11). The peak marked C, at $-0.08 \pm 0.05 \text{ mm s}^{-1}$, was due to iron cluster, because the isomer shift corresponded to that of iron cluster¹². In 10-resin, the peaks due to the magnetic splitting were weaker. In 30-resin, the peaks due to the quadrupole splitting were weaker, and the magnetic splittings were not detected. In 100-resin, only slight quadrupole splittings remained. That is to say, as the iron content in the resin decreases, the peaks due to the magnetic splitting and the quadrupole splitting successively weaken and disappear. The peaks due to iron cluster and Fe^{+} are present. In the inorganic glass-like carbon prepared by heat-treatment of the acetylferrocene-furfural resin, the liberated iron is in ultra-fine particles as shown in Fig. 1. The size of iron particles increases probably in the order of cluster, h.c.p. iron, distorted b.c.c. iron, and b.c.c. iron. In the resin heat-treated at 400°C , Fe^{+} ion is stabilised. In the present study, however, it could not be shown how the Fe^{+} ion is bonded chemically to the glass-like carbon matrix.

SEISHI YAJIMA
MAMORU OMORI

The Oarai Branch,
The Research Institute for Iron, Steel
and Other Metals,
Tohoku University,
Oarai, Ibaraki-ken, 311-13, Japan

Received 1 April; accepted 22 April 1977.

- ¹ Yajima, S., Hayashi, J., Omori, M. & Okamura, K. *Nature* **261**, 683 (1976).
- ² Yajima, S., Shishido, T. & Kayano, H. *Nature* **264**, 237 (1976).
- ³ Yajima, S., *et al.* *Nature* **264**, 238 (1976).
- ⁴ Yajima, S., Hayashi, J. & Okamura, K. *Nature* **266**, 521 (1977).
- ⁵ Yajima, S., Shishido, T. & Hamano, M. *Nature* **266**, 522 (1977).
- ⁶ Omori, M., Kuroki, M. & Yajima, S. *Bull. Chem. Soc. Japan* **48**, 1291 (1975).
- ⁷ Yajima, S. & Omori, M. *Chem. Lett.* **1972**, 843.
- ⁸ Yajima, S. & Omori, M. *Chem. Lett.* **1974**, 277.
- ⁹ Andreev, B. Ya., Dyagileva, L. M. & Feklisov, G. I. *Dokl. Akad. Nauk S.S.S.R.* **158**, 1348 (1964).
- ¹⁰ Jamieson, J. C. & Lawson, A. W. *J. appl. Phys.* **33**, 776 (1962).
- ¹¹ Walker, L. R., Wertheim, G. K. & Jaccarino, V. *Phys. Rev. Lett.* **6**, 98 (1961).
- ¹² McNab, T. K., Micklitz, H. & Barrett, P. H. *Phys. Rev. B*, **3**, 3787 (1971).

Bomb ^{14}C and human radiation burden

RADIOACTIVE fallout from nuclear weapon tests has resulted in a significant increase in mean radiation dose received by the world's population. While fission nuclides such as ^{137}Cs and ^{90}Sr are important sources of internal radioactivity, notable characteristics of artificial ^{14}C production are (1) its long half-life ($5,730 \pm 40 \text{ yr}$), (2) easy access via the food chain to all key molecules of body tissue, and (3) relatively long residence times in both stratosphere and troposphere of four and eight years respectively² (cf. mean residence time of particulate nuclear debris in the troposphere $\sim 30 \text{ d}$ ^{3a}). Although ^{14}C emits a low-energy β -particle ($\beta_{\text{max}} = 156 \text{ keV}$), its ease of incorporation in genetic material and the associated implications have provoked concern among several eminent scientists over artificial production of this radionuclide, even prior to the major testing period⁴⁻⁷. Recent data on levels of bomb ^{14}C and residence times of carbon in the human body are now available⁸. Thus excess human radiation burdens from this radioisotope are evaluated in this article under the assumption that no radiation dose, however small, can be regarded as entirely harmless biologically. This assumption, albeit a major one, is still adopted for radiation hazard considerations in the absence of conclusive evidence otherwise.

The absorbed dose, D_N , to soft tissue/bone marrow and bone-lining cells from decay of natural ^{14}C is $0.69 \text{ mrad yr}^{-1}$ and $0.74 \text{ mrad yr}^{-1}$ respectively. Since ^{14}C concentrations during the nuclear era are normally quoted as percentage

excess, $\Delta(\%)$, over the natural level, corresponding ^{14}C doses are directly expressed as $D_N(1 + \Delta/100)$. Annual absorbed doses from ^{14}C , both natural and man-made, are estimated for 1953–1973 (see Fig. 1) using $\tau = 6 \text{ yr}$ and $\tau = 30 \text{ yr}$ as representative mean values of residence times of carbon in soft tissue and bone collagen respectively⁸. Plotted values apply to mature tissue but may be used with minimal error for tissues under growth, with the exception of bone collagen. For this latter component, an approximate dose is estimated using conditions of linear increase in collagen with no carbon turnover. In this situation, the annual dose for a 20-yr old is simply the averaged annual dose contributions received before this age and reaches a maximum of $1.07 \text{ mrad yr}^{-1}$ for bone-lining cells, delivered in 1984 to individuals born in 1965.

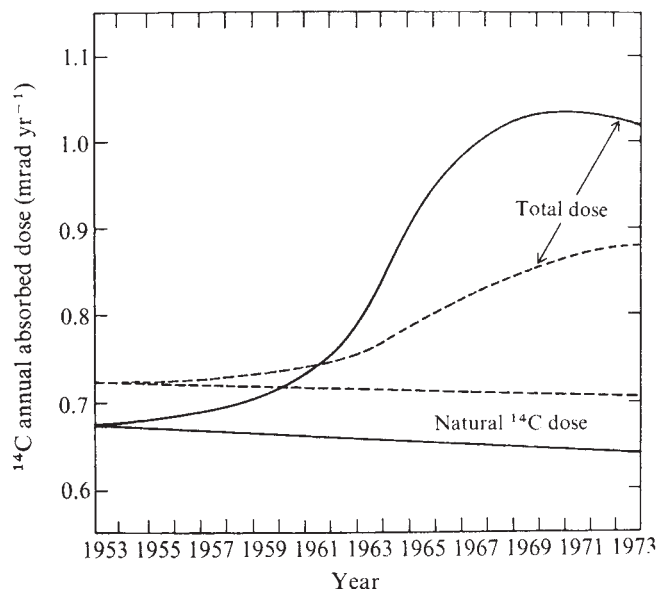


Fig. 1 Annual absorbed doses attributable to ^{14}C , 1953–1973. Total doses (natural + bomb ^{14}C) were calculated using tissue ^{14}C curves of Fig. 3, p.830; $\tau = 6 \text{ yr}$ and $\tau = 30 \text{ yr}$ were assigned as residence times to soft tissue/bone marrow and bone-lining cell carbon respectively. Natural ^{14}C dose rates were determined via theoretical consideration of the Suess effect¹⁰. ---, bone-lining cells; —, soft tissue/bone marrow.

Annual radiation doses for years 1975–2025 (see Fig. 2) are based on corresponding tissue ^{14}C levels. The latter data depend on future dietary ^{14}C content which has been estimated using (1) observed mathematical functions of atmospheric and oceanic ^{14}C levels⁹, and (2) a similar ocean water function but with an atmospheric ^{14}C activity decrease of 2% per yr, the present rate of decrease. The major uncertainty in Fig. 2 is associated with prediction of future atmospheric ^{14}C levels. With no further appreciable artificial production of this radioisotope, however, atmospheric ^{14}C content should reach pre-bomb levels by 2025, essentially via exchange of carbon with the oceans, and by dilution with inactive CO_2 (Suess effect). Again, natural ^{14}C data are included in Fig. 2, with estimates based on theoretical predictions of atmospheric $^{14}\text{CO}_2$ activities under the influence of the Suess effect alone¹⁰. Thus although environmental ^{14}C levels will approach pre-bomb concentrations by the first half of next century, net contributions from artificially produced ^{14}C are still appreciable.

For pertinent evaluation of possible biological effects of this radiation, total radiation received by any tissue over its critical years of function is more relevant than individual annual doses. Examples of cumulated doses to gonads (over 30 yr) and to bone marrow and bone-lining cells (over 60 yr)

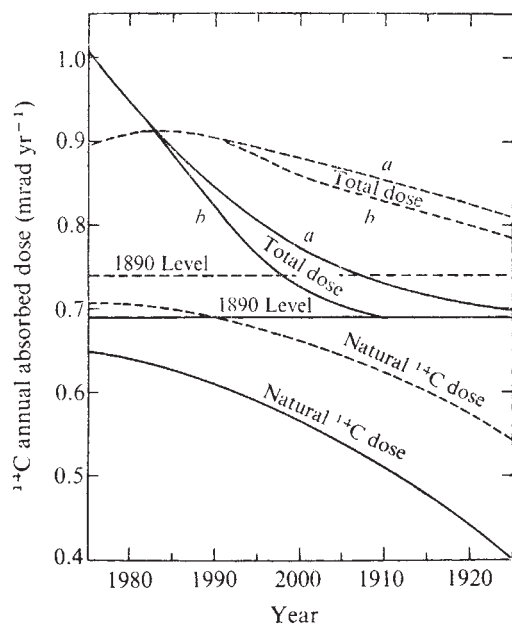


Fig. 2 Predicted annual absorbed doses attributable to ^{14}C , 1975–2025. Total doses to soft tissue/bone marrow and bone-lining cells were calculated from predicted tissue ^{14}C levels which are closely related to environmental ^{14}C activities. The latter data in line (a) were based on predicted future atmospheric and oceanic ^{14}C contents⁹, and in line (b) were based on an atmospheric decrease in ^{14}C content of $2\% \text{ yr}^{-1}$. Natural ^{14}C dose rates were assessed as in Fig. 1. Linear extrapolation of upper curves (a) indicates that the total absorbed doses to soft tissue/bone marrow and bone-lining cells will return to natural levels by years 2035 and 2050 respectively. ---, Bone-lining cells; —, soft tissue/bone marrow.

have been determined (see Table 1). Evaluation of the long-term genetic threat represented by any long-lived radioisotope utilises a quantity known as dose commitment, obtained *via* integration of dose over that time during which there is a significant contribution to the world population. For bomb ^{14}C , the dose commitment as calculated by UNSCEAR is 140 mrad and 170 mrad to soft tissue/bone marrow and bone-lining cells respectively³. These totals overestimate the potential radiation hazard from bomb ^{14}C , however, by total neglect of the Suess effect which, in the absence of artificial production of ^{14}C , would be expressed as a decrease in the tropospheric ^{14}C activity of $\sim 40\%$ below the natural level. Total accumulated doses in this article are obtained by addition of all annual absorbed doses in excess of that delivered by natural ^{14}C . Such a summation is performed until 2050 for bone-lining cells and to 2035 for soft tissue/bone marrow (see Fig. 2) to give values of 10 mrad and 8.5 mrad respectively. These totals, though well below ^{14}C dose commitments of the UNSCEAR report, are, we feel, more relevant for discussion of the real human radiation burden from bomb ^{14}C .

Table 1 Accumulated absorbed doses due to bomb ^{14}C

Period under consideration	Accumulated dose (mrad)	
	Soft tissue bone marrow	Bone-lining cells
pre 1954	—	—
1955–1984	7.26	3.16
1964–1993	8.81*	4.95
1970–1999	8.31	5.69
1985–2014	6.65	6.57*
1965–2024	bone growth during initial 20 yr	24.80* (60-yr dose)

*Maximum doses.

Assessment of the human radiation burden from bomb ^{14}C is the aim of this paper; thus it is important to relate the quantities determined here, which belong obviously in the category of low radiation doses to cellular modification, particularly chromosomal DNA damage. Possible radiation-induced chemical changes which occur in DNA have been reviewed recently¹¹ while Table 2 indicates the probability of damage per rad to DNA and other genetic components of a cell. Information in column 3, applicable to bomb ^{14}C , should be regarded with caution since data in column 2 are compiled from numerous research projects performed under a diversity of experimental conditions—dose rate, type of irradiation, but notably, type of cell¹². On the other hand, results of recent experiments using ^{14}C -methyl thymidine incorporated in mammalian DNA, substantiate at least the quoted extent of single strand breakage¹³. Allowance is made in Table 2 for repair of single strand breaks (up to 85% in mammalian DNA¹⁴) especially since such repair has also been shown to result in normal replication¹⁵. Unfortunately, interpretation of probabilities of cell damage due to bomb ^{14}C remains difficult since the lesion(s) responsible for the visible expression of biological damage has yet to be identified. For example, if each chromosome break is a critical lesion, the total absorbed dose due to bomb ^{14}C implies an ultimate probability of damage to any cell of 2×10^{-5} to 10^{-4} . Thus, given a birth rate of 3% in a world population of 2.5×10^9 , 1,500–7,500 offspring would possess altered genetic material resulting in a wide range of consequences.

Table 2 Estimates of biological damage to reproductive cells from ^{14}C β -irradiation (30-yr accumulated dose)

Description of damage to genetic material	Extent of damage (probability per cell)	
	Typical values per rad of ionising radiation	Contribution from bomb ^{14}C β -irradiation*
Single-strand breaks	10	0.01†
Double-strand breaks	1	0.006
Chromatid breaks	0.002–0.24	1.2×10^{-5} – 1.4×10^{-3}
Chromosome breaks	0.002–0.01	1.2×10^{-5} – 0.6×10^{-4}
Damaged bases	5–15	0.03–0.09

*Calculations performed using maximum bomb ^{14}C 30-yr absorbed dose.

†Probability based on 85% repair of single-strand breaks.

Local effects from the $^{14}\text{C} \rightarrow ^{14}\text{N}$ transition involve possible consequences of nuclear recoil, excited electronic state, and modified valence state of the daughter nucleus. In each respect, exact location of a ^{14}C atom is expected to be critical; and in certain situations, the transmutation effect may be the dominant damage factor. On the other hand, distinction between consequences of β -irradiation and local events is difficult for such a low-energy β -particle. Consequently there is conflicting experimental evidence for and against a positive transmutation role of ^{14}C decay¹⁶. Long-term experiments involving a ^{14}C general labelling technique have been completed recently by Lee and co-workers¹⁷; although genetic damage could be attributed to β -irradiation effects alone, the possibility of a transmutation effect associated with specifically-labelled DNA was not dismissed. With this background of inconclusive experimental evidence, the approach of Totter *et al.*⁷ remains the best theoretical treatment of a transmutation role during ^{14}C decay. Two situations for bomb ^{14}C are considered, (1) transmutation=mutation, cf. Totter *et al.*, and (2) transmutation at only one type of carbon site in DNA=mutation. Resultant values (see Table 3) represent extreme cases of a potential transmutation effect; once again, visible expression of this damage to cellular components cannot be predicted.

Table 3 Theoretical assessment of the transmutation effect of ^{14}C (applied to bomb ^{14}C)

total dose to gonads from bomb ^{14}C	≈ 10 mrad
time span involved	$= 80$ yr
thus mean annual dose to gonads (bomb ^{14}C)	$= 0.13$ mrad
cf. annual dose to gonads (natural ^{14}C)	$= 0.68$ mrad
dose rate due to bomb ^{14}C is	$\approx 20\%$ that due to natural ^{14}C
disintegration rate of natural ^{14}C	$= 13.56$ d.p.m. g^{-1} carbon
wt of DNA per cell	$\approx 8.6 \times 10^{-12}$ g
% of carbon in DNA	$\approx 37\%$
therefore no. of ^{14}C d.p.m.	$\approx 13.56 \times 8.6 \times 10^{-12} \times 0.37$ per DNA per cell
decays per DNA per cell per yr (natural ^{14}C)	$\approx 13.56 \times 60 \times 24 \times 365 \times 8.6 \times 10^{-12} \times 0.37$
	$\approx 2.27 \times 10^{-5}$ decays per cell per yr
decays per DNA per cell per yr (bomb ^{14}C)	$\approx \frac{2.27 \times 10^{-5}}{5}$ decays per cell per yr
	$\approx 4.5 \times 10^{-6}$ decays per cell per yr
Case 1	
transmutation = mutation, thus probability of mutation	
$= 4.5 \times 10^{-6}$ per cell per yr	
birth rate	$\approx 3\%$
world population	$\approx 2.5 \times 10^9$
total no. of yr	$= 80$
total offspring affected	$\approx 27,000$
Case 2	
transmutation at only one type of carbon site is responsible for a mutation, thus probability of mutation	
$= \frac{4.5 \times 10^{-6}}{40}$ per cell per yr*	
total offspring affected	≈ 700

*Approximately 40 different carbon sites in DNA.

Owing to their obvious limitations, the preceding approaches can only represent crude attempts at assessment of genetic damage due to bomb ^{14}C . For more relevant consideration of the social implications of artificial production of this radionuclide, reference is made to the latest UNSCEAR report which contains risk estimates of radiation-induced genetic disorders in humans^{3c}. These data are calculated with respect to damage to spermatogonia and oocytes, the germ cells most at risk, and are based on expected rates of damage at low doses and in conditions of chronic exposure, particularly applicable to ^{14}C . In the report, two approaches are employed to determine incidences of mutation of the zygotes—formed by union of male and female germ cells—per rad of low-dose irradiation administered to the parental generation. First, an incidence of damage of $\sim 1,500$ per 10^6 spermatogonia is obtained by extrapolation of genetic data involving mice to humans, taking into account the 20% increase in size of the human genome. The second value, 300 per 10^6 spermatogonia, is derived using a 3% spontaneous incidence in genetic damage among live born, together with a radiation doubling dose of 100 rads—that low-dose irradiation expected to double the natural mutation rate. Corresponding radiation damage to oocytes is considered negligible in comparison. Combination of these incidences of damage with 10 mrad, the total absorbed dose to the gonads due to bomb ^{14}C , yields upper and lower limits of 1,125 and 225 respectively for the number of individuals who will possess severe deleterious traits attributable to this radioisotope. It should be stressed that these figures have been derived using a linear dose-response relationship extrapolated to those levels of bomb ^{14}C , and should be assessed accordingly. Expression of this damage in first generation offspring is estimated at $\sim 20\text{--}55/5\text{--}10$ using an assumed range of dominance in man of $2\%\text{--}5\%^{3c}$. Consideration of minor deleterious effects have not been included in this computation but are expected to be more abundant even if no visible damage is apparent. This fact explains to some extent the lower values obtained here compared with those calculated in the preceding analyses.

With respect to late somatic effects, in particular an increase in number of leukaemia and cancer/tumour victims from effects of ^{14}C decay, there is a similar lack of data obtained at low doses and under chronic exposure conditions. Again, a linear dose-response relationship is assumed for radiation burden assessment although there is

evidence of a sigmoid relationship for low-density ionising radiation¹⁸. Spiers¹⁹ uses a figure of 100 cases per 10^6 individuals per rad as the total risk of leukaemia and different types of bone tumours; this induction rate is derived from the UNSCEAR report^{3d} and is rounded off owing to the many uncertainties involving dose-effect data. Thus this incidence of damage corresponds to a total of $\sim 2,000$ cases attributable to bomb ^{14}C ; again the role of transmutation of ^{14}C may be significant but this topic has been discussed recently by Tamers²⁰.

For both genetic and somatic considerations, the potential human radiation burden due to artificially produced ^{14}C and calculated using the assumptions stated is certainly significant and, therefore, cannot be dismissed. However, relative to the world population and in view of the time span over which this dose is delivered, some three to four generations, the conclusion of this study must be that such effects will occur unnoticed.

This work was generously supported by a grant from the Medical Research Council. M.J.S. thanks the National Environment Research Council for financial aid.

M. J. STENHOUSE*
M. S. BAXTER

Chemistry Department,
University of Glasgow,
Glasgow, UK

Received 27 May 1976; accepted 27 April 1977.

*Present address; Carbon-14/Tritium Measurements Laboratory, AERE Harwell, Oxfordshire, UK.

- Godwin, H. *Nature* **195**, 943–945 (1962).
- Walton, A., Ergin, M. & Harkness, D. D. *J. geophys. Res.* **75**, 3089–3098 (1970).
- Report of the United Nations Scientific Committee on the Effects of Atomic Radiation (United Nations, New York, 1972); 1, 39–40, 41–42, & 95; 2, 252–258, & 282; 2, 429.
- Pauling, L. *Science* **128**, 1183–1186 (1958).
- Sakharov, A. D. *Atomic Energy (U.S.S.R.)* (Eng. trans.) **4**, 757–762 (1958).
- Leipunsky, O. I. *Atomic Energy (U.S.S.R.)* (Eng. trans.) **3**, 1413–1425 (1958).
- Totter, J. R., Zelle, M. R. & Hollister, H. *Science* **128**, 1490–1495 (1958).
- Stenhouse, M. J. & Baxter, M. S. *Nature* **267**, 828–832 (1977).
- Gulliksen, S. & Nydal, R. *Proc. 8th Int. Radiocarbon Conf. New Zealand* **8**, C58–C72 (1972).
- Baxter, M. S. & Walton, A. *Proc. R. Soc. Lond.* **A318**, 213–230 (1970).
- Cerutti, P. A. *Die Naturwissenschaften* **61**, 51–59 (1974).
- Painter, R. B. *Current Topics in Radiation Research* **7**, 45–70 (North-Holland Publishing Co., Amsterdam 1970).
- Cleaver, J. E. & Burki, H. J. *Int. J. radn Biol.* **26**, 399–403 (1974).
- Ormerod, M. & Stevens, U. *Biochim. biophys. Acta* **232**, 72–82 (1971).
- Rasmussen, R. E., Reisner, B. L. & Painter, R. B. *Int. J. radn Biol.* **17**, 285–290 (1970).
- Krisch, R. E. & Zelle, M. R. *Adv. radn Biol.* **3**, 177–213 (1969).
- Lee, W. R., Segal, G. A. & Bensin, E. S. *Mutat. Res.* **16**, 195–201 (1972).
- Mays, C. W. *Proc. symp. Biolog. Environ. Effects Low-Level Radn.* **2**, 373–384 (IAEA, Vienna, 1976).
- Spiers, F. W. *3rd Int. Cong. Int. Radn. Prot. Ass., Washington DC* 439–447 (1973).
- Tamers, M. A. *Int. J. nucl. Med. Biol.* No. 1, 227–228 (1974).

Bomb ^{14}C as a biological tracer

THE release of ^{14}C in nuclear explosions, particularly those detonated in 1961–1962, has caused a major disturbance of the steady-state natural distribution of this radioisotope. The resultant pulse of bomb ^{14}C , detected initially in the atmosphere, has since been used extensively in evaluation of rates of exchange of carbon between major reservoirs of the dynamic cycle — notably atmospheric CO_2 and ocean water bicarbonate^{1–7}. Within the biosphere, however, investigation of human ^{14}C levels has been more limited despite the application of bomb ^{14}C as early as 1959 by Broecker and co-workers⁸ in the determination of residence times of carbon in several tissues. Perhaps the most specific result to emerge from initial research on this topic was the demonstration of almost total inertness of collagen in the adult human skeleton as evidenced by its failure to incorporate significant amounts of artificial ^{14}C (ref. 9). In other related ^{14}C measurements^{10–13}, a non-correspondence between contemporary human and atmospheric ^{14}C levels was often observed. Nydal and co-workers proposed a combination of three factors to account for such variations¹³ and concluded from a long term analysis of ^{14}C concentrations in human blood and hair samples that bomb ^{14}C enters the human body 1.4 years after production in the atmosphere, and remains there with a mean lifetime of 10 years. This article describes the continuation at Glasgow of an intensive investigation of the presence of bomb ^{14}C in humans, initiated by Harkness and Walton¹². Comparison by these workers of blood plasma protein ^{14}C data with predicted dietary ^{14}C levels¹⁴ has confirmed that a finite turnover of carbon in the human body does exist. Here, determination of ^{14}C activities of atmospheric CO_2 , dietary and human tissue samples enables a more detailed evaluation of residence times.

Atmospheric CO_2 was sampled monthly by adsorption in KOH solution (4 M) at three UK locations (see Fig. 1). Representative samples from each location were analysed for ^{14}C content since ^{14}C measurement of all samples was impracticable. Human tissues, normally about 20 g, were obtained unfixed from postmortem examinations. After washing, the tissues were homogenised and freeze-dried and subsequently samples were combusted directly to CO_2 or pretreated, before combustion, with Bloor's solvent (ethanol–diethyl ether, 3:1 v/v) to provide protein- and lipid-rich components for analysis.

Grossly diseased aortas from victims of atherosclerosis were forwarded from New Orleans, USA. An additional artery sample was obtained at Glasgow. Although the former samples had been stored in formalin, it was felt that this source of carbon contamination would be eliminated during lipid extraction. After washing, the intima was carefully stripped from the other two arterial layers; plaques were separated from relatively normal tissue and extracted overnight at room temperature with chloroform–methanol (2:1 v/v). The extract yielded on evaporation an oil from which lipids were selectively removed into ether–hexane (1:1 v/v). This solution was washed with warm saline solution, dried with Na_2SO_4 , and evaporated to dryness. A stream of nitrogen was passed over each residue for several hours to ensure total removal of solvent. Samples obtained in this way weighed typically 1–2 g.

Biological samples were combusted in a stream of oxygen and the resultant CO_2 absorbed in KOH solution. Details are given elsewhere¹⁵ of the subsequent purification of CO_2 necessary for gas-proportional anti coincidence counting. ^{14}C activities were measured in a 2.61 counter, filled to a pressure corresponding to one atmosphere at 15 °C; under these conditions the background count rate was 4.9 c.p.m. with an observed deviation of -0.01 c.p.m. per mbar rise in atmospheric pressure.

Tables 1 and 2 include the ^{14}C results of all human tissues analysed, after correction for isotopic fractionation by analysis of the $^{13}\text{C}/^{12}\text{C}$ ratio of each sample relative to the P.D.B. Chicago Limestone Standard. Corresponding theoretically

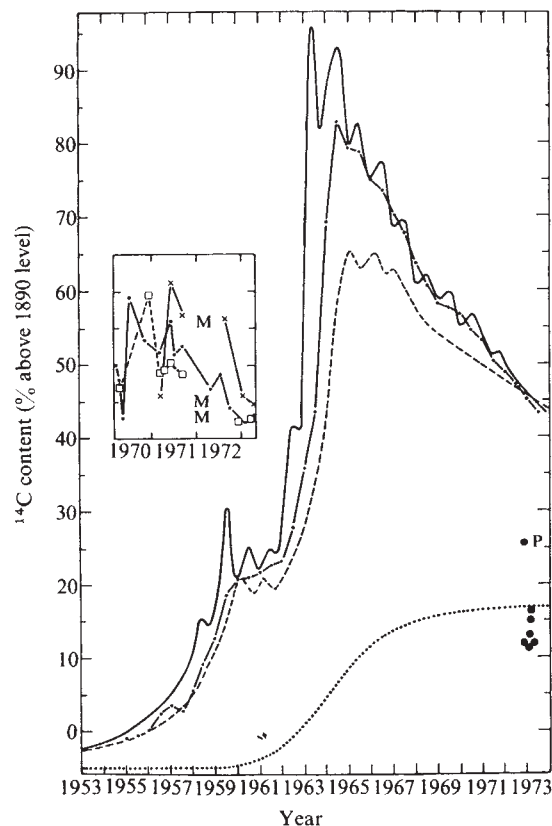


Fig. 1 Environmental ^{14}C concentrations, 1953–1973. ●, Fish samples obtained fresh in March 1973 (●P = plaice); M, meat obtained fresh in January 1972. —, N. Hemisphere troposphere; ---, S. Hemisphere troposphere., Mixed ocean layer; — · — · —, computed dietary level. Inset: $^{14}\text{CO}_2$ activities at UK latitudes, —, Lerwick (60°08'N, 01°11'W); —□—, Snowdon (53°03'N, 04°00'W); —×—, Chilton (51°31'N, 01°20'W).

determined values were evaluated by application of a suitable box-model treatment of carbon turnover in tissues. The potential of bomb ^{14}C in this system results from nature's unique labelling mechanism whereby ^{14}C incorporation in biological material yields a uniform distribution throughout each molecule—no particular molecular sites are favoured. ^{14}C , both natural and man-made, is assimilated by plants by photosynthesis with subsequent transfer to humans through the food chain. Thus the ^{14}C content of any tissue may be regarded as a function of dietary ^{14}C level and of the residence time of carbon in that tissue. From this basic model (Fig. 2a) it can be shown that, for adult tissue, in which an essentially stationary state exists (carbon input = output), the ^{14}C content at any time is given by the equation (see a similar expression by Jansen¹⁸)

$$\Delta_T^{ni} = \Delta_T^0 + \sum_{n=1}^i \left\{ \Delta_D^{ni} - \Delta_T^{(n-1)i} \right\} \left\{ 1 - \exp(-k_T i) \right\} \quad (1)$$

where T and D refer to tissue and diet respectively; $\tau = 1/k_T$, the mean residence/turnover time, is the average time required by a tissue for replacement of an amount of carbon equivalent to the tissue itself. Thus for a given τ value, and with 'i' fixed at 0.5 yr, tissue ^{14}C concentration is dependent only on the cumulative differences between the previous ^{14}C activity and the mean ^{14}C content of the diet over the subsequent half-yearly interval. By appropriate choice of zero time (t_0 before the effects of nuclear weapon tests) a plot of Δ_T^{ni} against 'ni' indicates build-up of bomb ^{14}C in any tissue.

Table 1 ^{14}C content of human tissues*

Donor/date of death†	Liver	Kidney	Spleen	Heart	Muscle	Lung	Brain	Miscellaneous
M/21: 9/71	54.6±0.9	51.2±1.6	46.0±0.8	42.4±1.4	55.9±1.3	53.9±1.6	—	Pancreas: 52.2±1.1 Testes: 56.0±1.3 Thyroid: 43.3±1.6
F/21: 9/71	55.2±1.0	53.1±1.2	54.8±1.6	54.5±1.3	51.7±1.3	51.1±1.1	52.3±1.1	Pancreas: 54.6±1.1 Ovaries: 50.6±2.0 Thyroid: 50.7±1.1
M/37: 4/72	48.4±0.9	—	—	47.3±1.0	—	—	—	Testes: 49.8±1.8
M/39: 8/71	42.3±1.2	—	38.6±1.0	39.8±1.2	—	—	48.1±0.9	Thyroid: 37.9±1.5 Pancreas: 46.0±1.2
M/42: 9/71	47.7±1.0	33.9±0.9	43.5±2.5	51.3±1.5	—	37.8±1.0	51.9±1.0	Thyroid: 43.3±0.9 Testes: 37.8±1.0
M/50: 1/71	43.3±0.8	44.6±0.8	—	52.3±1.9	50.2±1.3	—	—	—
M/61: 10/72	48.5±1.0	45.3±0.8	43.4±0.9	50.4±0.8	—	—	—	—
	P45.0±1.1	P43.7±1.0	P46.2±1.0	P44.9±1.0	—	—	—	—
M/64: 11/72	47.5±0.8	46.3±0.9	43.2±1.0	—	—	44.0±1.2	—	—
	P48.7±0.9	P48.4±1.0	P42.8±0.9	P40.9±0.9	—	P46.5±2.2	—	—
	L47.2±1.1	—	—	—	—	—	—	—
F/64: 4/71	—	49.5±1.0	45.6±1.2	41.5±1.5	48.2±1.0	—	49.1±0.9	Adipose tissue: 46.9±1.2 Ovaries: 47.5±1.5
M/67: 1/73	49.4±0.8	46.6±0.8	—	46.8±0.8	—	—	—	—
	P49.4±1.0	P48.5±0.9	—	P45.4±0.8	—	—	—	—
	L48.7±1.2	—	—	L51.1±0.8	—	—	—	—
M/72: 3/71	53.2±1.0	55.2±0.9	51.5±1.0	50.7±1.0	51.9±1.1	51.2±1.0	50.4±1.0	Pancreas: 54.7±0.9 Testes: 49.8±1.0 Bone(mf): 48.9±0.8 Bone(col): 10.4±2.7
F/87: 1/73	48.0±0.9	46.2±0.9	—	—	—	—	—	—
	P49.1±1.0	P47.9±0.8	P51.1±0.9	—	—	—	—	—
M/5: 9/71	52.8±0.9	50.4±1.5	44.9±0.8	—	52.5±1.0	—	50.5±0.9	Pancreas: 53.5±1.2

*These data are described in greater detail in refs 15 and 16.

†Age and sex of donor, month and year of death are given.

P, protein-rich; L, lipid-rich; mf, marrow fat; col, collagen; muscle analysed, forearm.

See also footnotes to Table 2.

In practice, dietary ^{14}C content has not been measured with sufficient frequency over the last 20 yr but may be related to two more adequately monitored systems—atmospheric CO_2 and ocean carbonate/bicarbonate. Many such correlations are well established: plant life and cereals are good indicators of ambient atmospheric ^{14}C activities^{19,20}. In addition, the food types included in Fig. 1 reflect, fairly closely, environments of their origin, although recognition must be given to existence of carbon turnover in animals as well as in humans. Further relationships used in this study have been derived from previous environmental ^{14}C measurements²¹ coupled with data on stock/production of foodstuffs; the latter allows assessment of lag times between food growth and consumption (see Table 3).

With this knowledge, a comprehensive survey of dietary habits of the UK population was performed to evaluate

percentages of protein, lipid and carbohydrate supplied by different commodities. Note was made of the hemisphere of origin of food imports since marked disequilibrium between north and south hemisphere tropospheric CO_2 activities existed until 1971. The mean ^{14}C content of a typical UK diet

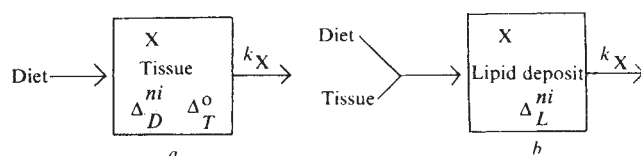


Fig. 2 Single compartment models of carbon turnover. X turns over with rate constant k_X , that fraction of X replaced in unit time (1 yr). X represents in (a), tissue carbon, supplied by a dietary source only; and in (b), arterial lipid carbon, supplied by the blood by material of dietary and endogenous origins. $\Delta_T^0 = ^{14}\text{C}$ content of tissue at t_0 , computed zero time; $\Delta_T^{ni} = ^{14}\text{C}$ content of tissue ni years after t_0 ; i is a suitable time increment—0.5 yr; Δ_D^{ni} = mean ^{14}C content of diet over the n^{th} i period after t_0 . For the non-stationary system: N^{ni} = no. of carbon atoms ni yr after t_0 (tissue/lipid); N_T^0 = no. of carbon atoms at t_0 —at birth/11 yr of age.

at any time (see Fig. 1) is given by the expression

$$\Delta_D^{ni} = 0.24\Delta_{TN}^{ni} + 0.19\Delta_{TN}^{(n-1)i} + 0.16\Delta_{TN}^{(n-2)i} + 0.28\Delta_{TS}^J + 0.02\Delta_{TS}^{ni} + 0.035\Delta_{TS}^{(n-2)i} + 0.035\Delta_{TS}^J + 0.04\Delta_M^{ni}$$

TN and TS signify north and south hemisphere tropospheres respectively, while M denotes the mixed surface-ocean layer; $(n-1)i$ and $(n-2)i$ denote lag-time effects and J refers to the preceding period of maximum photosynthesis.

For tissues under growth, non-stationary conditions necessitate a modified approach. Use is made here of the general shape of growth curve obtained for most tissues²²; this curve, to a first approximation, may be regarded as a combination

Table 2 ^{14}C content of human tissue

Donor	Sample	$\Delta(\%) \pm 1\sigma$
M/66: 3/73	Liver (normal)	50.0±0.9
	Liver (cancered)	55.2±1.4
M/63: 2/73	Spleen (normal)	51.0±0.9
	Liver (cancered)	53.0±1.2
M/59: 2/73	Kidney (cancered)	48.1±0.8
M/68: 1963	Intimal lipid	4.8±0.6
M/67: 1964	Intimal lipid	9.7±0.9
M/69: 1964	Intimal lipid	17.9±0.8
F/86: 6/73	Intimal lipid	47.1±0.9
	Adventitia	31.6±0.8

*Normally, sample activities are quoted initially as $\delta^{14}\text{C}$, the per mil enrichment over $0.95 \times \text{NBS Oxalic Acid Standard activity}$ (=1890 wood). Here, after correction for isotopic fractionation using the Lamont VIII formula¹⁷, final values, expressed as $\Delta(\%)$, are shown in column 3. ^{14}C concentrations defined on this scale can be regarded as the percent deviation from pre-bomb or natural levels; $\Delta = 0\%$ corresponds to a natural ^{14}C concentration, and $\Delta = 50\%$ to a concentration $1.5 \times$ natural level. Errors, arising from uncertainties in ^{14}C activity measurement only, are quoted to one standard deviation ($\pm 1\sigma$).

Table 3 Dietary information

Main food categories	Contribution to typical diet (%)	Environmental relationship, Δ_D^{ni}
Milk/dairy products	23.0	Δ_T^{ni}
Cereals/flour/sugar	31.5	Δ_T^{ni}
Meat/poultry	34.0	$\Delta_T^{(n-1)i}, \Delta_T^{(n-2)i}$
Fruit/vegetables—fresh	3.4	Δ_T^{ni}
—other	3.7	$\Delta^{(n-1)i}$
Fish	4.4	Δ_M^{ni}

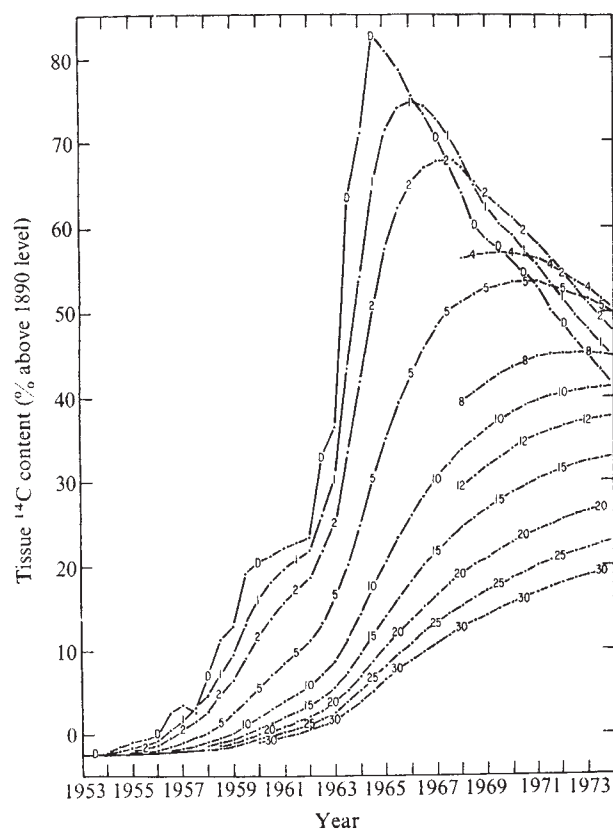
of two linear growth periods: from birth to 11 yr, and from 11 to 20 yr. In these conditions the comparable equation is

$$\Delta_T^{ni} = \Delta_T^0 + \sum_{n=1}^n \left\{ 1 - \frac{N_T^{0+} (n-1)(N_T^i - N_T^0)}{N_T^{ni}} \exp(-k_T i) \right\} \times \left\{ \Delta_D^{ni} - \Delta_T^{(n-1)i} \right\} \quad (2)$$

Since the first 2 yr after birth are best expressed by an exponential growth function, N_T^0 , rather than the absolute number of carbon atoms at birth, is that value obtained by extrapolation to birth (t_0) of the linear section of growth. Both N_T^0 and $N_T^i - N_T^0$, the increment in number of carbon atoms per 'i' time period, may be expressed as fractions of the mature tissue. For the 5-yr-old tissue series, errors involved in basic assumptions of the non-stationary system are considered too great to justify such a treatment. Nevertheless, ^{14}C activities of this series, with the exception of one organ, reflect dietary ^{14}C levels of the 18-month period prior to analysis. The different dietary pattern of a child may be responsible for the observed low ^{14}C content of spleen.

Model treatment of lipid accumulation in arteries also

Fig. 3 Predicted ^{14}C content of mature tissues, 1953–1973. Each curve represents a different mean residence time of tissue carbon. Steady-state conditions and first-order kinetics of carbon turnover are assumed to exist. —●—; $\tau = i$ yr; —D—; $\tau = 0$.



involves a non-stationary system. Numerous pathological findings indicate a linear incidence in number of atherosclerotic lesions with time from 20–25 years of age²³. Thus on the basis of a corresponding increase in amount of carbon deposited, the corresponding version of equation (2) is:

$$\Delta_L^{ni} = \Delta_L^{(n-1)i} + \sum_{n=2}^n \left\{ 1 - \frac{(n-1)}{n} \exp(-k_L i) \right\} \times \left\{ \Delta_L^{ni} - \Delta_S^{(n-1)i} \right\} \quad (3)$$

where L and S refer to lipid and source carbon respectively; the latter is taken to be of composition similar to that of

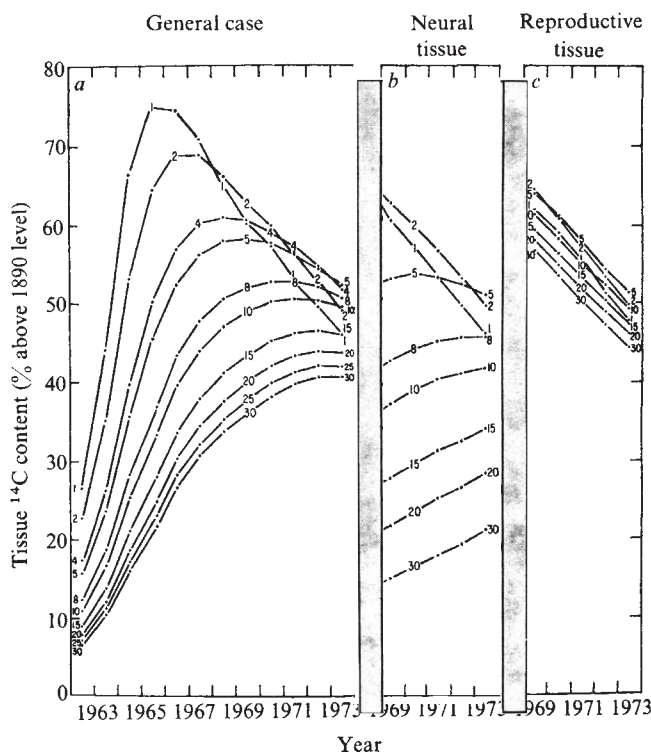


Fig. 4 Predicted ^{14}C content of tissues under growth, 1953–1973. Non steady-state conditions exist: (a) The general series of curves was based on two linear growth periods—birth–11 yr, and 11–20 yr. (b) For brain tissue, curves were computed for linear growth from birth–5 yr; thereafter steady-state conditions were applied. (c) For reproductive tissue, equation (1) was applied from birth to 15 yr, followed by linear growth from 15–20 yr. —●—; $\tau = i$ yr.

corresponding lipids in blood—30% dietary and 70% endogenous origins²⁴. Although most tissues are capable of preparation of cholesterol and its esters, two major components of the arterial extract, liver and gastrointestinal tract are considered to be main sites for biosynthesis of these fractions²⁵. Thus each Δ_S term is subdivided as follows:

$$\Delta_S^{ni} = 0.3\Delta_{D2}^{ni} + 0.7\Delta_T^{ni}$$

Δ_T^{ni} data were obtained from estimated ^{14}C activities of a relevant tissue such as liver, while dietary emphasis was placed on those constituents high in cholesterol, total fat, and saturated fatty acid content; in this case

$$\Delta_{D2}^{ni} = 0.48\Delta_{TN}^{ni} + 0.09\Delta_{TN}^{(n-1)i} + 0.28\Delta_{TN}^{(n-2)i} + 0.095\Delta_{TS}^{ni} + 0.035\Delta_{TS}^{(n-2)i} + 0.02\Delta_M^{ni} \quad (4)$$

Figures 3–5 contain series of curves computed from Δ_D^{ni} input relevant to each model, each curve corresponding to a

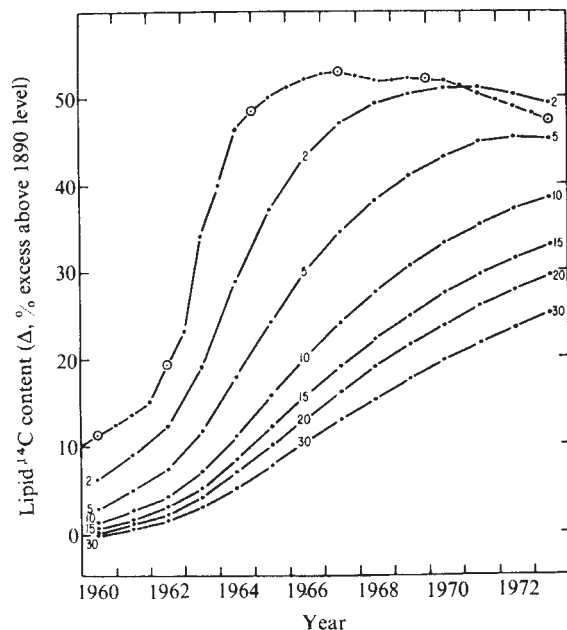


Fig. 5 Predicted ^{14}C content of arterial lipid deposits, 1953–1973. A similar situation to Fig. 4 exists, but with a different source of carbon input (30% dietary, 70% endogenous). No deposit was present at zero time, t_0 . —●—: $\tau = i$ yr; —○—: $\tau = 0$.

different mean residence time of tissue carbon. Comparison of the contents of Table 1 with the relevant figure yielded estimates of the turnover time, τ years, of carbon in that tissue. Uncertainties in this comparison are related to (1) ^{14}C activity measurement ($\pm 1\sigma$) and (2) diet–atmosphere relationship (estimated at ± 1 yr). Figure 3 shows that, for τ values 0–5, uncertainty in the diet–atmosphere relationships is predominant, while for longer residence times, both (1) and (2) are equally significant. Thus sensitivity of the bomb ^{14}C tracer method is expressed in terms of years rather than months. At this level of resolution, differences in turnover rates of normal against cancerous tissue are undetectable, certainly at the time of sampling. Similarly, for the purposes of this study, protein, lipid and whole tissue carbon exhibit no significant τ value variation. On the other hand, the long residence time obtained for bone collagen confirms previous findings^{9,14} of comparative inertness of skeletal carbon relative to soft tissue.

Significant variations in τ values exist, not only in tissues of

one donor, but also between the same tissues in different individuals, with no apparent trends as a function of age. Of the soft tissues, only thyroid carbon can be assigned, with confidence, a residence time ≥ 9 yr; elsewhere, in many cases, choices of τ values/range of τ values exist. This ambiguity does not imply that both selections are possible but merely signifies the inability of the model to distinguish between certain residence times owing to the proximity, at the time of analysis, of certain curves. However, it is difficult to reconcile the smaller of two choices of τ with the number of tissues ($> 25\%$) in which a residence time ≥ 8 yr is indicated. Comparison of Figs 3 and 4 shows that, for $\tau = 0/1$, there is negligible difference between corresponding Δ_T curves; yet about 35% of both 21-yr-old tissue series have ^{14}C content higher than either corresponding data from contemporary donors, or Δ_D values at the period of sampling. Considered together, these tissue results imply a residence time > 0 , but typically 2–5 yr. Thus as shown in Table 4, preference is given to upper τ ranges. In view of limitations imposed on the model at this time, a representative value of the mean turnover time of soft tissue carbon is taken as 6 ± 4 yr. This figure must be regarded as a maximum estimate which includes recycling of carbon within a tissue as well as interaction between different organs. Obviously the advantage of the bomb ^{14}C approach lies in detection of slow turnovers, typically > 6 yr; results can then be used to complement short term tracer studies which often fail to establish accurate upper limits on turnover rates. It should be stressed that τ values determined here are representative of each tissue in its entirety, or at least of its protein–lipid fractions. Extreme variations in metabolic rates must exist within tissues then, if results of this study are considered in conjunction with previous experimental observations which indicate much faster turnover rates of tissue amino acids^{26,27}.

Further application of bomb ^{14}C as a biological tracer is illustrated by Table 5 which implies a residence time, for carbon in intimal lipid deposits, of at least 5 yr. In the context of the model used, however, $\tau > 10$ yr seems more probable, provided that lipid is deposited directly from the blood. Another possible mechanism of lipid accumulation—*in situ* production—has been proposed²⁸, particularly for phospholipid build-up; accordingly, low Δ_L values may reflect ^{14}C activities of the endogenous precursor(s) of deposited carbon. In this respect, the ^{14}C content of adventitia (outer arterial layer) indicates the presence of about 25% natural ^{14}C , probably attributable to collagen²⁹. These data, while insufficient for detailed interpretation, encourage further measurements along a similar theme to determine the possible pathogenesis or timescale of treatment of atherosclerosis.

Table 4 Most probable residence times (τ , yr) of tissue carbon

Donor	Liver	Kidney	Spleen	Heart	Muscle	Lung	Brain	Testes/ Ovaries	Pancreas	Thyroid
M/21	1–6	8–10	15	20–30	1–6	5–8	—	1–10	7–9	15–25
F/21	1–6	6–8	1–7	1–7	8–10	8–10	5–6	10–15	1–7	9–10
M/37	7	—	—	7	—	—	—	5–6	—	—
M/39	9	—	11	10–11	—	—	7	—	7–8	11–12
M/42	7	13–14	8–10	5–6	—	11–12	5–6	—	—	9
M/50	8–9	8	—	5–6	6	—	—	10–12	—	—
M/61	6–7	8	9	6	—	—	—	—	—	—
M/64	P 8	P 9	P 7–8	P 8–9	—	8–9	—	—	—	—
	7	7–8	9	—	—	—	—	—	—	—
	P 6–7	P 6–7	P 9–10	P 10–11	—	P 7–8	—	—	—	—
F/64	L 6–7	—	—	—	—	—	—	—	—	—
	—	6–7	8	9–10	7	—	6–7	7	—	—
	6–7	7–8	—	7–8	—	—	—	—	—	—
M/67	P 6–7	P 6–7	—	P 8	—	—	—	—	—	—
	L 6–7	—	—	L 5–6	—	—	—	—	—	—
	5	1–4	6	6	5–6	6	6	6	4–5	—
M/72	7	7–8	—	—	—	—	—	—	—	—
F/87	P 6–7	P 7	P 5–6	—	—	—	—	—	—	—

M/72 Bone collagen, $\tau \geq 30$ yr; marrow fat, $\tau = 6$.
M/63 Spleen (normal), $\tau = 5–6$; Liver (cancerous), $\tau = 2–4$.

M/66 Liver (normal), $\tau = 5–6$; cancerous = 3–4.
M/59 Kidney (cancerous), $\tau = 6–7$.

Table 5 Mean residence times (τ) of arterial carbon

Donor	Sample	τ (yr)
M/68	Lipid extract	10-30
M/67	Lipid extract	5-25
M/69	Lipid extract	2-10
F/86	Lipid extract	0-5
	adventitia	15

This work was generously supported by a grant from the Medical Research Council. M.J.S. thanks the National Environment Research Council for financial aid. Assistance of the many suppliers of samples is gratefully appreciated; in particular the contribution of Professor W. A. Harland who, besides collecting tissue samples, provided constructive comment, is acknowledged. Technical aid was provided by A. Miller, while F. Williams and H. M. Blauer derived the necessary computer programmes.

M. J. STENHOUSE*
M. S. BAXTER

Chemistry Department,
University of Glasgow,
Glasgow, UK

Received 27 May 1976; accepted 27 April 1977.

*Present address: Carbon-14/Tritium Measurements Laboratory, AERE Harwell, Oxfordshire, UK.

1. Bien, G. & Suess, H. E. *Radioactive Dating and Methods of Low-level Counting* 105-115 (I.A.E.A., Vienna, 1967).
2. Münnich, K. O. & Roether, W. *Radioactive Dating and Methods of Low-level Counting* 93-104 (I.A.E.A., Vienna, 1967).
3. Nydal, R. *Radioactive Dating and Methods of Low-level Counting* 119-128 (I.A.E.A., Vienna, 1967).
4. Young, J. A. & Fairhall, A. W. *J. geophys. Res.* 73, 1185-1200 (1968).
5. Walton, A., Ergin, M. & Harkness, D. D. *J. geophys. Res.* 75, 3089-3098 (1970).
6. Rafter, T. A. & O'Brien, B. J. *Proc. 8th Int. Radiocarbon Conf. New Zealand* 8, C17-C37 (1972).
7. Gulliksen, S. & Nydal, R. *Proc. 8th Int. Radiocarbon Conf. New Zealand* 8, C58-C72 (1972).
8. Broecker, W. S., Schulert, A. & Olson, E. A. *Science* 130, 331-332 (1959).
9. Libby, W. F., Berger, R., Mead, J. F., Alexander, G. V. & Ross, J. F. *Science* 146, 1170-1172 (1964).
10. Nydal, R. & Louseth, K. *Nature* 206, 1029-1031 (1965).
11. L'Orange, R. & Zimen, K. *Naturwissen* 55, 492 (1968).
12. Harkness, D. D. & Walton, A. *Nature* 223, 1216-1218 (1971).
13. Nydal, R. Lövseth, K. & Syrtstad, O. *Nature* 232, 418-421 (1971).
14. Harkness, D. D. & Walton, A. *Nature* 240, 302 (1972).
15. Stenhouse, M. J. & Baxter, M. S. *Radiocarbon* 18, 161-171 (1976).
16. Farmer, J. G., Stenhouse, M. J. & Baxter, M. S. *Radiocarbon* 14, 326-330 (1972).
17. Broecker, W. S. & Olson, E. A. *Radiocarbon* 3, 176-204 (1961).
18. Jansen, H. S. *Proc. 8th Int. Radiocarbon Conf. New Zealand* 8, F1-F26 (1972).
19. Baxter, M. S. & Walton, A. *Proc. R. Soc. Lond. A* 321, 105-127 (1971).
20. Tauber, H. *Radiocarbon* 9, 246-256 (1967).
21. Stenhouse, M. J. thesis, Univ. Glasgow (1975).
22. Spiers, F. W. *Radiosotopes in the Human Body* (Academic, New York and London, 1968).
23. Eggen, D. A. & Solberg, L. A. *Lab. Invest.* 18, 571-579 (1968).
24. Kaplan, J. A., Cox, G. E. & Taylor, C. B. *Arch. Pathol.* 76, 359-368 (1963).
25. Ho, K. J., Taylor, C. B. & Biss, K. in *Atherosclerosis, Proc. 2nd Int. Symp. Chicago* 271-273 (1969).
26. Schoenheimer, R. *The Dynamic State of Body Constituents* (Harvard University Press, Cambridge, Mass 1942).
27. Neuberger, A. & Richards, F. F. in *Mammalian Protein Metabolism*, 1 243-296 (Academic, New York and London, 1964).
28. Zilvermit, D. B. in *Atherosclerosis, Proc. 2nd Int. Symp. Chicago* 35-41 (1969).
29. Kao, K. Y., Hilker, D. M. & McGavack, T. H. *Proc. Soc. expt. Biol. Med.* 106, 121-124 (1961).

Lunar cycle and timing of mating season in Serengeti wildebeest

WILDEBEEST (*Connochaetes taurinus* Thomas) herds numbering nearly 1 million (refs 1, 2) on the Serengeti Plains, Tanzania, produce about 80% of their young during a 3-week period in January or February³. This is short compared with the 2-month birth period of the wildebeest in South Africa⁴, or to the 3-4-month birth period of buffalo and other grazers in the Serengeti⁵. A suggested advantage for the closely synchronised wildebeest birth period is that it reduces the probability of hyaena (*Crocuta crocuta* Erxleben) predation on the young^{3,6-8}. I report here evidence which suggests that the lunar cycle is the external timer for the synchronisation.

I have calculated individual conception dates from records of (1) foetus weights collected randomly and (2)

birth dates. Conception date can be calculated from foetus weight (ref. 9) (Fig. 1), and from birth date with a known gestation length. This lies between 8 and 9 months (refs 10, 11) and I have assumed 255 d.

Figure 1 shows the frequency distribution of conception dates from a random sample of 33 foetuses collected in 1971. Also shown are conception dates for 1964 and 1965, and the mating period for 1962 and 1963, all of which were calculated from information given by Watson¹¹. The distribution of conception dates for 1964, 1965 and 1971, together with other evidence³, show that conceptions are synchronised into one discrete time period with a relatively sharp onset. I have called this the 'mating period'.

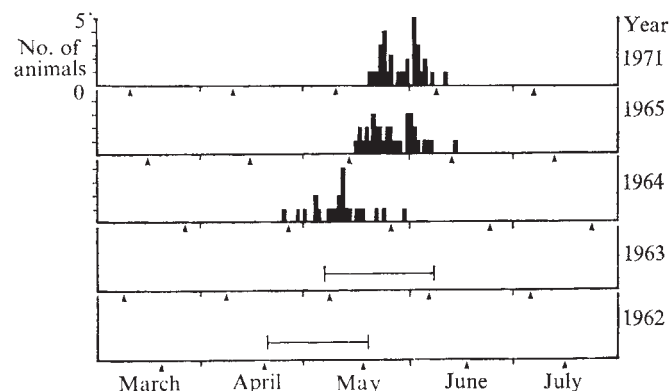


Fig. 1 Distribution of wildebeest conception dates in different years. Full moon is indicated by arrows. Conception dates were obtained from the relation between foetus age in days (t) and its weight in grams (W): $W^{1/3} = a(t - t_0)$, where t_0 is 20% of gestation length (t_g) and equals 51 d. 'a' is a growth constant and equals $W_g^{1/3}(t_g - t_0) = 0.143$. W_g is the birth weight of 25 kg. For the few small foetuses under 100 g, age was calculated directly from growth curves extrapolated from cattle¹³.

In Fig. 1 the times of full moon are shown by arrows. The mating period, where adequately observed (1962-1965, 1971), falls between the full moons. It starts approximately at the full moon and covers the new moon period. If there was no association between the timing of the mating period and the occurrence of full moon, the latter would be expected in the middle of this mating period on some occasions. To test for such an association I compared the starting date of the mating period with the date of the nearest full moon (Table 1), for there is additional information on the starting date derived from calving in the years 1968 and 1972 in a population 50 km to the east in Ngorongoro Crater, and also in the Serengeti in 1973 (A. and J. Root, personal communication). Estes¹ observed the calving in three discrete herds in Ngorongoro in 1973.

If there was no relation between starting date and full moon then these dates should fall randomly within 15 d

Table 1 Absolute difference between full moon and start of the mating period

Place	Year	Start of mating period	Full moon	Difference between start and full moon (d)
Serengeti	1962	20 April	19 April	1
Serengeti	1963	7 May	8 May	1
Serengeti	1964	25 April	26 April	1
Serengeti	1965	16 May	14 May	2
Ngorongoro	1967	21 May	23 May	2
Ngorongoro	1971	15 May	10 May	5
Serengeti	1971	20 May	10 May	10
Ngorongoro a	1972	25-27 April	28 April	3
b	1972	25-27 April	28 April	3
c	1972	25-27 April	28 April	3
Serengeti	1972	27 April	28 April	1

either side of any full moon. Table 1 shows that 10 of the 11 dates fall within 7 d and 9 fall within 5 d of the full moon, strongly suggesting a relationship (the latter figures have $\chi^2=7.25$ with Yates correction, d.f.=1, $P<0.01$).

Supporting evidence for this relationship comes from the dates when first calves were observed by the park warden M. Turner¹² in the years 1956–1965. Conception dates have been calculated from these (Table 2). If there was no relationship with the moon, then these dates should fall at any time during the 30 d since the preceding full moon. In fact in eight out of ten years the first calculated conception fell within the first 10 d following the full moon, and all fell within the first 15 d ($P<0.01$, Kolmogorov–Smirnov test¹³). Wildebeest can delay birth for a short time⁶ but not to the extent that it could have produced the above observations.

Table 2 Conception dates calculated from first observed calving dates

First observed calving date	Conception date	Time since full moon (d)
11 November 1956	29 February 1956	4
4 December 1957	25 March 1957	10
16 December 1958	5 April 1958	2
6 December 1959	27 March 1959	3
2 December 1960	23 March 1960	11
20 December 1961	9 April 1961	9
12 November 1962	1 March 1962	9
27 October 1963	13 February 1963	5
23 November 1964	12 March 1964	14
28 December 1965	17 April 1965	2

All conception dates have been calculated using the mean gestation length of 255 d. This may be 5–10 d shorter or longer. Consequently until gestation time is more accurately known, one cannot say which lunar phase is important, but only that some aspect of it, as yet undetermined, is influencing conception.

Figure 1 shows that conceptions take place at approximately the same time each year. For conceptions to occur only in the lunar cycles of April and May an external variable showing an annual rhythm must also be influencing the animals. Although at this latitude (3°S) day-length alters by only 20 min through the year and is changing at its slowest in May (Fig. 2), it is possible that a combination of photoperiod and lunar cycle may act to trigger mating.

Alternatively local climatic conditions could act as a trigger for they are changing markedly at the start of mating in April (Fig. 2). In particular, rainfall and the amount of green grass upon which the animals feed¹⁴ decline rapidly at this time. This hypothesis is illustrated in Fig. 3. When rainfall or some related factor such as

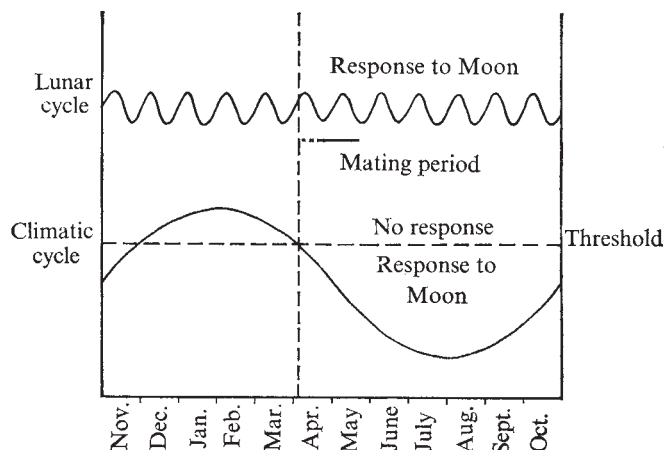


Fig. 3 Suggested model for the synchronisation of conceptions in wildebeest. When the climatic cycle (rain, food) falls below a threshold the animals become sensitive to the lunar cycle which switches on ovulation or allows conception to occur.

nutritional intake falls below a certain threshold, then the animals may become sensitive to changes in the lunar cycle. This still leaves a choice of two or three cycles to choose from. It is possible that if all females conceive in the first lunar cycle after the beginning of the decline in rainfall, then no conceptions would be observed in the following potentially suitable cycles. Indeed the birth data of Estes³ in Ngorongoro show that some but not all small herds mate over a 2-month period, the distribution of conceptions being bimodal, with each peak coincident with each new moon. The data indicate that if the full moon occurs in the first half of April then conceptions are delayed by one more lunar cycle. These hypotheses can be tested by making predictions on the timing of the mating and birth seasons for future years.

To my knowledge this is the first case of a strong relationship between conception and the lunar cycle in mammals. There may be a similar effect in Australian water-buffalo (*Bubalus bubalis*)¹⁴ and perhaps a small effect on man¹⁵. For other mammals the data remain inconclusive¹⁶. The effect of lunar cycles may still be of relevance to the timing of breeding seasons in tropical regions, however.

I thank J. H. Myers, C. J. Krebs, R. M. F. S. Sadleir, B. C. R. Bertram and N. Gilbert for their help. H. Kruuk, G. B. Schaller, and A. and J. Root kindly provided unpublished records.

A. R. E. SINCLAIR

Serengeti Research Institute,
P.O. Seronera, Arusha, Tanzania

Received 7 February; accepted 22 April 1977.

*Present address: Institute of Animal Resource Ecology, University of British Columbia, Vancouver, British Columbia, Canada.

- Sinclair, A. R. E., *E. Afr. Wildl. J.* 11, 93–107 (1973).
- Norton-Griffiths, M. E. *Afr. Wildl. J.* 11, 135–149 (1973).
- Estes, R. D. *E. Afr. Wildl. J.* 14, 135–152 (1976).
- Fairall, N. *Zool. Afr.* 3, 189–210 (1968).
- Sinclair, A. R. E. *E. Afr. Wildl. J.* 12, 169–183 (1974).
- Estes, R. D. *Nature* 212, 999–1000 (1966).
- Kruuk, H. *The Spotted Hyena* (Chicago University Press, Chicago, 1972).
- Wilson, E. O. *Sociobiology* (Belknap Press, Cambridge, Massachusetts, 1975).
- Hugget, A. St. G. & Widdas, W. F. *J. Physiol. Lond.* 114, 306–317 (1951).
- Talbot, L. M. & Talbot, M. H. *Wildl. Monogr.*, Chestertown No. 12 (1963).
- Watson, R. M. *J. Reprod. Fert.*, Suppl. 6, 287–310 (1969).
- Turner, M. I. M. *Chief Park Wardens Report 1965–1966, Serengeti National Park* (mimeographed, 1966).
- Siegel, S. *Nonparametric Statistics* (McGraw-Hill, New York, 1956).
- Sinclair, A. R. E. *The African Buffalo* (Chicago University Press, Chicago, 1977).
- Menaker, W. & Menaker, A. *Am. J. Obst. Gynec.* 77, 905–914 (1959).
- Sadler, R. M. F. S. *The Ecology of Reproduction in Wild and Domestic Animals* (Methuen, London, 1969).

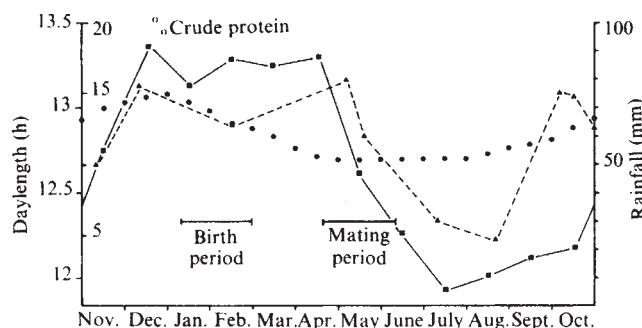


Fig. 2 The birth and mating season of Serengeti wildebeest in relation to daylength (●), mean monthly rainfall 1963–1972 (■), and % crude protein in the diet in 1971 (▲).

Passage of *Trypanosoma brucei rhodesiense* through the peritrophic membrane of *Glossina morsitans morsitans*

THE developmental cycle of African sleeping sickness trypanosomes in tsetse flies (*Glossina* spp.) as described by Robertson¹ has recently been questioned², especially the route whereby trypanosomes get from within the mid-gut of the fly to the salivary glands, where they develop into forms capable of infecting fresh mammalian hosts. We have shown² that trypanosomes developing in the mid-gut of *Glossina morsitans morsitans*, which are usually regarded as possessing low penetrative powers, are able to penetrate and pass through the gut cells of the fly. We present here more evidence of the penetrative powers of the fly mid-gut forms of *Trypanosoma brucei rhodesiense*—penetration of the peritrophic membrane of *G. m. morsitans* before their penetration of the gut cells.

An essential part of the developmental cycle of *T. brucei* spp. trypanosomes in the fly is the passage of the trypanosomes from the endoperitrophic space of the midgut to the ectoperitrophic space. The traditionally accepted route for this passage¹ is for the trypanosomes to escape from the confines of the peritrophic membrane around its often ragged posterior ends, where the mid- and hindgut join, and so gain access to the ectoperitrophic space. It has been suggested that trypanosomes *en route* to the salivary glands have to re-enter the endoperitrophic space by penetrating

the peritrophic membrane in the region of the proventriculus. Here the membrane is freshly secreted and hence soft^{3,4}, and electron micrographs of trypanosomes penetrating the peritrophic membrane in this region have been published⁵. This portion of the peritrophic membrane has also been suggested as a possible site for the initial access of the trypanosomes to the ectoperitrophic space^{6,7}. The penetration we describe, however, takes place in the middle section of the mid-gut, halfway between the proventriculus and the junction of the mid- and hindgut.

The techniques for the preparation of infected blood meals, the feeding, details of infecting trypanosomes and dissection of the flies, together with those for fixation and EM examination, were as used previously². Throughout our studies of this model, flies not sacrificed continued to harbour trypanosomes, and a satisfactory proportion developed mature salivary-gland infections and transmitted the infection to new mammalian hosts.

We found that as digestion of the infected blood meal proceeds and the volume of the material within the peritrophic sac diminishes, the peritrophic membrane becomes thrown into folds, and pockets are formed which trap the trypanosomes (Fig. 1a). The peritrophic membrane probably has several layers, the most easily recognisable of these being an inner osmophilic layer and an outer, paler, wider and more diffuse one. These layers separate (Fig. 1b) and the trypanosomes penetrate the osmophilic layer and gain entry to the space between the separated layers of the membrane (Fig. 1c). Penetration takes place most commonly at the apex of the membrane folds, which is the site

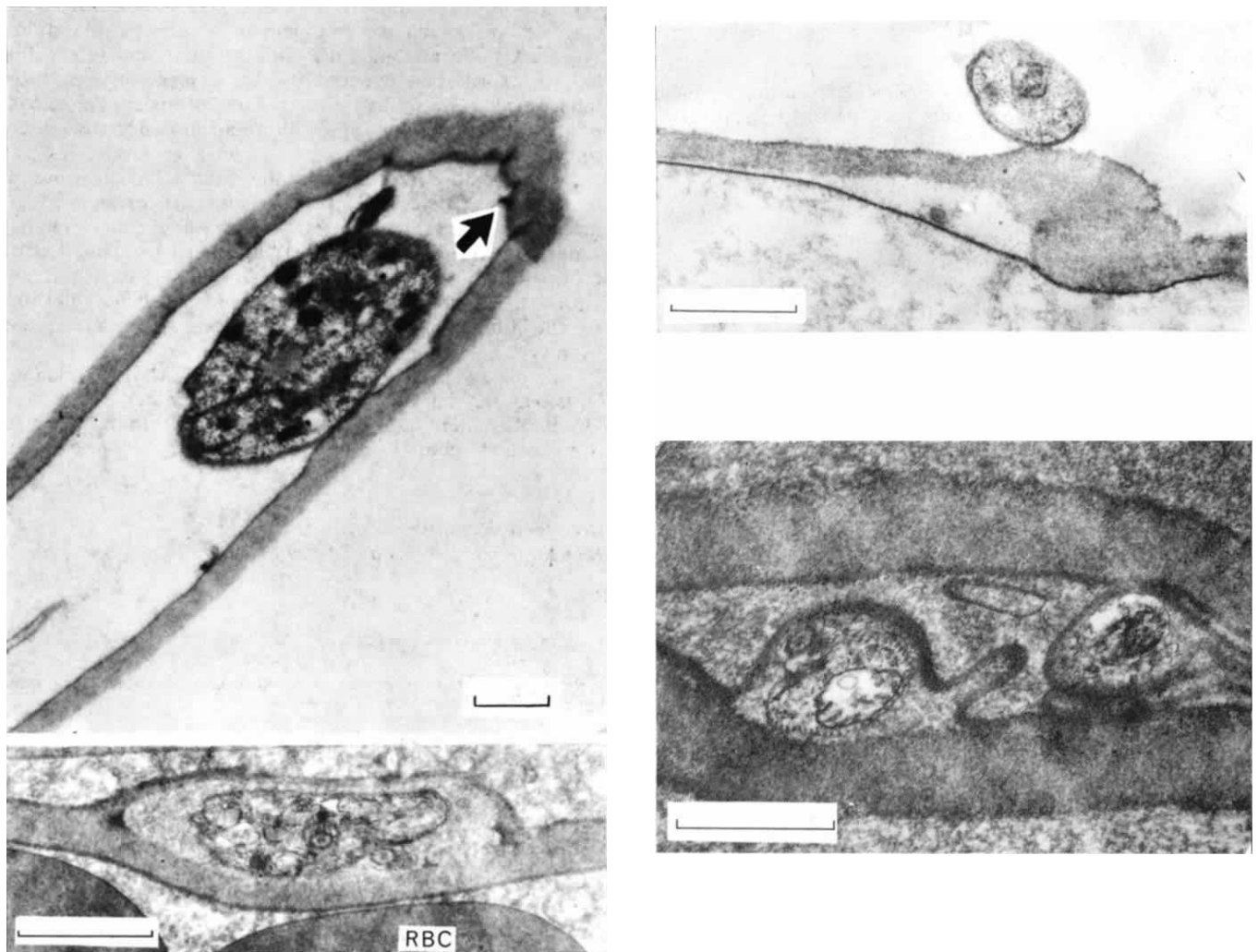


Fig. 1a A trypanosome trapped in a fold of the peritrophic membrane; note the 'buckling' of the inner layer at the apex (arrowed). b, Separation of the peritrophic membrane layers. c, Two trypanosomes lying between layers of the membrane in a fold. d, A trypanosome lying within the outer thicker layer. The cells of the blood meal (RBC) are below the peritrophic membrane. Bars represent 1 μ m.

where the separation of the layers is most frequently seen, particularly where the osmophilic layer is on the inside of the curve. Figure 1d shows trypanosomes lying within the thick layer of the peritrophic membrane.

Penetration of the peritrophic membrane occurred throughout the central two-thirds of the midgut, and its frequency was related to the trypanosome concentration, which was greatest in these central sections. We found no evidence of penetration occurring in the hindmost portion of the midgut, and the few trypanosomes found at this level appeared moribund. Nor did we find trypanosomes involved in the peritrophic membrane anterior to this midgut region. If any trypanosomes initially entered the ectoperitrophic space through the soft membrane in the proventricular region³ they had disappeared within 70 h of the infected feed, and we found no evidence of any trypanosomes outside the membrane until 10 d had elapsed.

Nine days after an infected feed all the trypanosomes were still within the peritrophic sac; by 11 days most of the trypanosomes had passed to the ectoperitrophic space, only 5% remaining within the sac, and a similar proportion were still involved in active penetration. The passage of successive waves of trypanosomes across the membrane from additional infected feeds took less time to complete—12 d compared with 14 d for a single infected feed.

The peritrophic membrane was found to be separated into two layers only in the presence of trypanosomes. Digestive enzymes may have a role in this process, but difference in rigidity between the layers of the membrane could certainly account for the separation seen at the sharp folds which occur when the sac contents decrease during the digestion of a blood meal. Lehan⁸ has shown in *Stomoxys calcitrans* that the thin layer of the peritrophic membrane contains no chitin (a rigid material) while the underlying layer does.

Epidemiological studies of the vectors of African sleeping sickness have traditionally relied on examination of the tsetse fly salivary glands for evidence of infection. In areas of endemic disease, the infection rates recorded by this method often fail to reach 0.8% (ref. 9). Any alternative life-cycle pathways of trypanosomes within the vector might suggest additional sites to examine within the fly when looking in the field for potential transmitters of African trypanosomiasis. Are we at present overlooking a large number of tsetse flies already capable of transmitting disease by only examining the salivary glands? Are the salivary glands continuously primed with trypanosomes coming from other parts of the fly (the haemocoel for instance) so as to maintain the population of metacyclic trypanosomes; or, once primed, is this sufficient to maintain salivary gland infection for the rest of the life of the fly? These are fundamental questions which a detailed study of all stages of the life-cycle of the trypanosomes in the fly would resolve.

This work was supported by the Ministry of Overseas Development. We thank Miss Susan Stamford, Mr Mark Holland and Mr David Day for technical assistance.

D. S. ELLIS
D. A. EVANS

Department of Medical Protozoology,
London School of Hygiene and Tropical Medicine,
Keppel Street, London WC1, UK

Received 3 March; accepted 4 May 1977.

- ¹ Robertson, M. *Phil. Trans. R. Soc. B* 203, 161–184 (1913).
- ² Evans, D. A. & Ellis, D. S. *Nature* 258, 231–233 (1975).
- ³ York, W., Murgatroyd, F. & Hawking, F., *Ann. trop. Med. Parasit.* 27, 347 (1933).
- ⁴ Fairbairn, H. *Ann. trop. Med. Parasit.* 52, 18 (1958).
- ⁵ Steiger, R. F. *Acta tropica* 30, 64–168 (1973).
- ⁶ Willet, K. C. *Expl. Parasit.* 18, 290–295 (1966).
- ⁷ Freeman, J. C. *Acta tropica* 30, 347 (1973).
- ⁸ Lehan, M. J. *J. Insect Physiol.* 22, 1551–1557 (1976).
- ⁹ Buxton, P. A. *Mem. Lond. Sch. Hyg. trop. Med.* 10, 619–623 (1955).

Rhodopsin content and electroretinographic sensitivity in light-damaged rat retina

EXPOSURE of rats to continuous illumination leads to cellular deterioration within the retina and a concurrent reduction of electroretinogram (ERG) responses^{1,2}. Grignolo *et al.*³ showed that at early stages, this 'damage' due to light is hypertrophic in that tubules and vesiculae appear in receptor cell outer segments. They surmised that losses in retinal functioning are initially due to disorganisation rather than degeneration. But, experiments on recovery from light damage² indicate that ERG sensitivity returns as the rhodopsin-containing disks of the outer segments are regenerated even though their original lamellar structure is not restored. It is thereby implied that as long as visual pigment is present at functional sites, the normal structure of the outer segments may not be crucial to ERG responsiveness. Here we describe an experiment which quantitatively examines the role of visual pigment remaining in the albino rat retina which has been exposed to continuous low-intensity light. The ERG sensitivity, rhodopsin content and morphological state of the retina were assessed for various exposure durations. Our findings indicate that reductions in visual pigment levels due to constant light persist indefinitely in the dark. Furthermore, the log ERG (b wave) sensitivity varied directly with the rhodopsin content of exposed retinas even as progressive deterioration was apparent. With respect to this log-linear relationship, light damage resembles 'normal' light adaptation due to short term bleaching exposure.

An especially interesting feature of our results is that even extensive damage could be effected by lights which were considerably dimmer than any previously reported. This lighting was provided by 'cool white' fluorescent bulbs with a sheet of translucent, white plexiglass interposed to give a diffusely-illuminated environment of 50–55 lx measured at right angles to the walls of the enclosures. This light level corresponds to that which is found in many indoor environments. Ambient temperature of the cages was constant at $24 \pm 1^\circ\text{C}$ and the rats moved freely within them.

Constant light exposures ranged from 24 to 108 h and were followed by a period of 12 h in darkness. ERGs were then recorded using a Ag-AgCl cotton wick electrode positioned on the cornea of rats which were anaesthetised

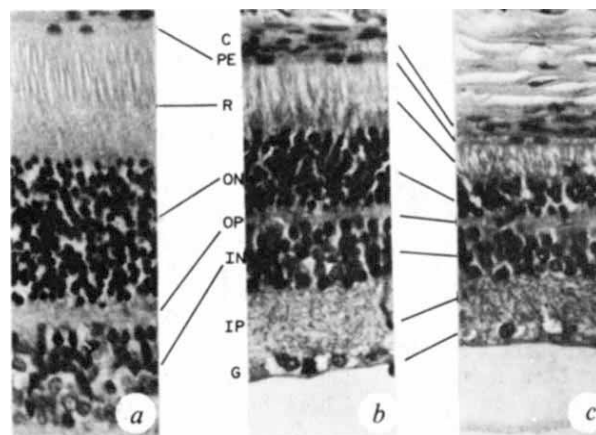


Fig. 1 Light micrographs of transverse sections through albino rat retinas showing different degrees of light damage. a, Control retina; b, 60-h exposed retina; c, 108-h exposed retina. Tissue was fixed in 4% glutaraldehyde and stained with haematoxylin-eosin. Sections were taken from a central region of the retina. C, Choroid; PE, pigment epithelium; R, receptor inner and outer segments; ON, outer nuclear layer; OP, outer plexiform layer; IN, inner nuclear layer; IP, inner plexiform layer; G, ganglion cells. IP and G extend beyond the lower edge of the micrograph of the control retina (a). $\times 800$.

with sodium pentobarbital. The indifferent electrode was placed on the unstimulated eye and shielded from stray light. Threshold intensity was determined for a criterion b wave of 25–30 μ V using a 20-ms monochromatic flash of 500 nm. Following electrophysiological measurements, the retinas were excised and used for either histological study or rhodopsin assay in which pigment was extracted with 15% Triton X-100.

Control rats were taken directly from an animal room with a 12-h light, 12-h dark cycle of very dim light. After 12-h dark adaptation, their threshold intensities were measured and equated to 1.0 on a relative scale so that log threshold is equal to 0.0. Rhodopsin content of damaged rats is expressed in this study as a percentage of that found in the controls.

Figure 1 shows the degenerative changes of the rat retina resulting from increasing durations of exposure to constant, low intensity light. With 60-h exposure, a shortening and disorganisation of the outer segment layer can be observed (Fig. 1b). Several cell bodies of the outer nuclear layer were missing and compaction of the retina was evident. No obvious changes in the pigment epithelium were revealed by light microscopy. Rats exposed for 108 h were almost completely lacking in identifiable outer segments (Fig. 1c). The outer nuclear layer was reduced to less than one-third

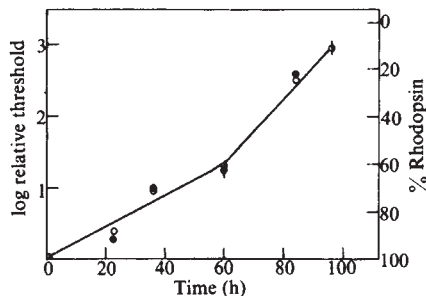


Fig. 2 Time course of sensitivity and rhodopsin content changes with constant exposure to low-intensity illumination. ○, Rhodopsin content; ●, log relative threshold. Standard deviation bars are drawn upon averaged log threshold data points for groups of four animals.

of its original thickness reflecting the loss of many cell bodies; but, of those remaining many seemed to have intact inner segments. In addition, considerably fewer cell nuclei had degenerated in the peripheral retina as compared with the central regions. This distribution of damage has also been observed by previous investigators^{3,4}.

The level of rhodopsin in the retina, as measured after the 12-h dark adaptation period, was continuously lowered with time of exposure to low intensity light. Ordinarily complete regeneration of bleached visual pigment occurs within 2–3 h in the dark⁵. Only 24 h of constant light exposure was sufficient to cause a prolonged reduction in the amount of rhodopsin as compared to the control value. With 108-h exposure, only 11% of the rhodopsin remained. In Fig. 2 it is seen that a steady rise in log b-wave threshold paralleled the depletion of visual pigment. Figure 3 shows the relationship between log threshold and rhodopsin content of the retina for various exposure durations. A linear function between these two variable is obtained which has also been found for retinas bleached (in a conventional manner) with relatively short light exposures^{3,6} and for visual pigment levels reduced by vitamin A deprivation⁷. Thus, our findings demonstrate a third way in which rhodopsin can be varied so that the log ERG threshold depends directly on visual pigment concentration in the retina. It should be noted that the observed rise in threshold due to a reduction in the rhodopsin level is disproportionately higher than what would be expected simply on the basis of the quantal catching ability of the photoreceptors (Fig. 3).

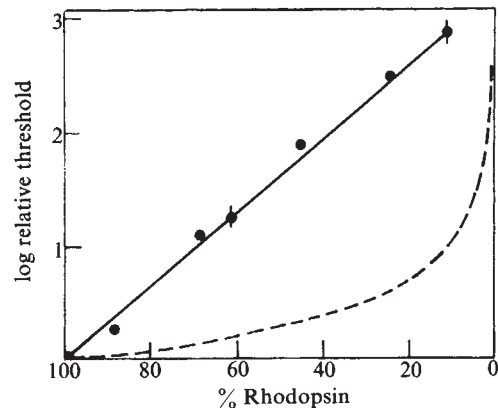


Fig. 3 Relationship between log ERG threshold and reduced visual pigment levels due to constant light exposure. Vertical bars denote the standard deviations of groups of four animals. The dotted line represents the rise in log threshold expected if it were only affected by changes in the probability of quantal absorption by the photoreceptors.

In view of histological changes that are produced by low intensity light exposures, our result implies that even seriously distorted rod outer segments are capable of initiating retinal signals as long as some rhodopsin is present within them. If the cells which generate the b wave can still function at or near their normal capacity following a damaging light exposure, they might only suffer a loss of input in a way that doesn't differ much from 'bleaching' light adaptation. The Müller cells of the inner retina are a likely candidate since they have been implicated as the source cells of the b wave⁸ and have been shown to be spared even after severely damaging light exposures⁴.

This work was supported by the US NSF and NIH. We thank Dave Watson for technical assistance.

LAURENCE M. RAPP
THEODORE P. WILLIAMS

Department of Biological Science,
Florida State University,
Tallahassee, Florida 32306

Received 3 March; accepted 26 April 1977.

- Noell, W. K., Walker, V. S., Kang, B. S. & Berman, S. *Invest. Ophthalmol.* 5, 450–473 (1966).
- Kuwabara, T. *Am. J. Ophthalmol.* 70, 187–198 (1970).
- Grignolo, A., Orzalesi, N., Castellazzo, R. & Vittone, P. *Ophthalmologica* 157, 43–49 (1969).
- Kuwabara, T. & Gorn, R. A. *Arch. Ophthalmol.* 79, 69–78 (1968).
- Dowling, J. E. *J. gen. Physiol.* 46, 1287–1301 (1963).
- Weinstein, G. W., Hobson, R. R. & Dowling, J. E. *Nature* 215, 134–138 (1967).
- Dowling, J. E. & Wald, G. *Proc. natn. Acad. Sci. U.S.A.* 46, 587–608 (1960).
- Miller, R. F. & Dowling, J. E. *J. Neurophysiol.* 33, 323–341 (1970).

Species specific agglutination of eggs by bindin isolated from sea urchin sperm

SPERMATOZOA adhere to eggs during sea urchin fertilisation^{1,2}. The adhesion is between the acrosome process of the sperm and the vitelline layer covering of the egg³. During the sperm acrosome reaction, the membrane of the extending acrosome process becomes coated with protein derived from the sperm acrosome vesicle^{3,6}. We believe^{4,5}, like others^{3,6}, that the acrosome vesicle protein binds the sperm to the egg by interacting with glycoprotein receptors on the egg vitelline layer^{4,5,7,8}. We have isolated the insoluble contents of the acrosome vesicle and have shown it to be composed of a single protein of molecular weight 30,500 which we have named bindin⁴. In most interspecies inseminations sea urchin sperm fail to adhere to the vitelline layers of foreign eggs and fertilisation does not occur, even though sperm undergo the acrosome reaction in the presence of the foreign eggs⁹. Presumably this failure of gamete adhesion occurs because bindin, on the sperm acrosome process, does not bind to the glycoprotein receptors on the vitelline layers of eggs of another species. We report here the species specific agglutination of unfertilised eggs by isolated bindin. This demonstrates the

preservation of the species specific recognition event between the isolated protein and its receptor which should facilitate the *in vitro* reconstruction and analysis of the molecular mechanism of gamete adhesion.

Gametes were spawned into seawater by injecting 0.5 M KCl into adults. Egg jelly layers were dissolved by a 2-min exposure to pH 5 seawater followed by return to pH 8 seawater. Eggs exhibited 100% normal fertilisation as judged by the elevation of the fertilisation envelope within 60 s after a test insemination. Eggs suspensions used in

agglutination tests were stored at 0 °C no longer than 2 h before being discarded. Insoluble, particulate bindin was isolated from sperm as previously described¹. After the 13,000g centrifugation the pellet of bindin was transferred to siliconised 30 ml correx glass centrifuge tubes and sonicated with the minimum intensity to achieve complete suspension. The bindin was sedimented three times and resuspended by sonication in 15 ml fresh calcium free seawater pH 5.8 (ref. 4). The final pellet was resuspended in normal seawater, the protein concentration was determined¹⁰ and storage was at 0 °C until use. Eggs were placed in wells of an agglutination titre plate, various amounts of bindin suspension added and the plate agitated by either a reciprocal or a rotary motion at 20 °C.

Figure 1 shows that agglutination of eggs by bindin is species specific. The final concentration of bindin in the agglutination mixture (Fig. 1) was $430 \mu\text{g ml}^{-1}$ for *Strongylocentrotus purpuratus* and $230 \mu\text{g ml}^{-1}$ for *S. franciscanus*. By serially diluting the bindin we determined that for both species $3.5 \mu\text{g ml}^{-1}$ was the minimum bindin concentration needed to agglutinate homologous eggs to the extent shown in Fig. 1. In other words, the concentration of bindin used in this experiment far exceeds the minimum needed to achieve maximum agglutination. That the two bindins do not cross react at high concentrations with eggs of the other species suggests the recognition by bindin of its receptor on the egg surface must be unique to each species. Some loss of specificity occurred as the eggs aged, however. *S. purpuratus* eggs older than 2 h post-spawning sometimes show slight agglutination by *S. franciscanus* bindin, but this was never true for the reciprocal cross mixture. If bindin is partially solubilised in 8 mM urea, 2% mercaptoethanol, then dialysed back into seawater, it still agglutinates eggs but the species specificity is lost. The native structure of bindin seems necessary for the species specificity of agglutination. We have reported⁴ that formaldehyde and glutaraldehyde fixed eggs are agglutinated by bindin. The species specificity of agglutination of aldehyde-fixed eggs is not reliably reproducible for reasons we have not investigated.

The older literature on fertilisation¹¹⁻¹³ stressed that the species specificity of fertilisation is based on the existence of specific 'fertilisins' in the jelly layer surrounding the egg. This hypothesis was recently questioned and re-evaluated by Collins¹⁴. We found that the addition of solubilised egg jelly has no effect on the agglutination of eggs by bindin, and that the species specific attachment of sperm to sea urchin eggs is based on the interaction of a sperm protein, bindin, with the egg vitelline layer. Sperm attachment to the zona pellucida covering of the mammalian egg is also a species specific phenomenon^{15,16} which might involve a similar type of recognition between mammalian sperm bindin and zona bindin receptors. Our work with sea urchins is the first demonstration of the species specific interaction of a protein isolated from sperm with the surface of an animal egg.

This work was supported by the US National Institute of Child Health and Human Development.

CHARLES G. GLABE
VICTOR D. VACQUIER

Department of Zoology,
University of California,
Davis, California 95616

Received 18 March; accepted 2 May 1977.

- ¹ Tegner, M. J. & Epel, D. J. *exp. Zool.* **197**, 31-58 (1976).
- ² Vacquier, V. D. & Payne, J. E. *Exp. Cell Res.* **82**, 227-235 (1973).
- ³ Summers, R. G., Hylander, B. L., Colwin, L. H. & Colwin, A. L. *Am. Zool.* **15**, 523-551 (1975).
- ⁴ Vacquier, V. D. & Moy, G. W. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- ⁵ Glabe, C. G. & Vacquier, V. D. *J. Cell Biol.* submitted.
- ⁶ Dan, J. C. in *Fertilization 1* (eds Metz, C. B. & Monroy, A.) 237-293 (Academic, New York, 1967).
- ⁷ Aketa, K. *Embryologia* **9**, 238-245 (1967).
- ⁸ Schmell, E., Earles, B. J., Breaux, C. & Lennarz, W. J. *J. Cell Biol.* **72**, 35-46 (1977).

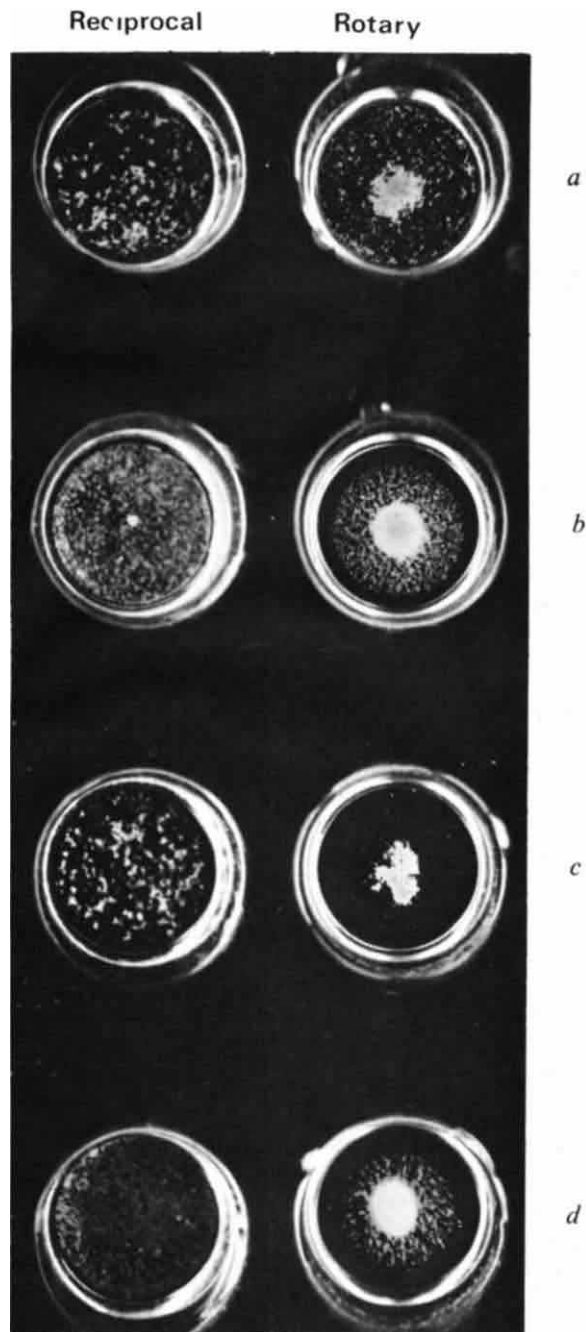


Fig. 1 The agglutination of *Strongylocentrotus purpuratus* and *S. franciscanus* eggs by bindin is species specific. 0.25 ml *S. purpuratus* bindin ($850 \mu\text{g ml}^{-1}$) or *S. franciscanus* bindin ($450 \mu\text{g ml}^{-1}$) was placed in wells (bottom diameter 15 mm) of a Linbro FB-54 agglutination titre plate. 0.25 ml 2% v/v suspension of eggs was added and the plate mixed by either reciprocal or rotary motion with minimum agitation to keep the eggs in suspension. After 2-5 min the agglutination patterns were photographed. Agglutination is easy to discern in homologous mixtures of bindin and eggs. *a*, *S. purpuratus* eggs + *S. purpuratus* bindin = agglutination. *b*, *S. purpuratus* eggs + *S. franciscanus* bindin = no agglutination. *c*, *S. franciscanus* eggs + *S. franciscanus* bindin = agglutination. *d*, *S. franciscanus* eggs + *S. purpuratus* bindin = no agglutination.

- ⁹ Summers, R. G. & Hylander, B. L. *Exp. Cell Res.* 100, 190-194 (1976).
¹⁰ Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* 193, 265-275 (1951).
¹¹ Lillie, F. R. *J. exp. Zool.* 16, 523-590 (1914).
¹² Tyler, A. & Tyler, B. S. in *Physiology of Echinodermata* (ed. Boolootian, R. A.) 683-741 (Wiley, New York, 1966).
¹³ Monroy, A. *Chemistry and Physiology of Fertilization* (Rinehart Holt and Winston, New York, 1965).
¹⁴ Collins, F. *Dev. Biol.* 49, 381-394 (1976).
¹⁵ Hartman, J. F., Gwatkin, R. B. L. & Hutchison, C. F. *Proc. natn. Acad. Sci. U.S.A.* 69, 2767-2771 (1972).
¹⁶ Hartman, J. F. & Hutchison, C. F. *J. Reprod. Fert.* 37, 61-66 (1974).

Expression of α -galactosidase in preimplantation mouse embryos

THE most widely accepted model for X-chromosome dosage compensation in female mammals is the single-active X hypothesis (Lyon hypothesis)¹. X-Chromosome dosage compensation probably occurs at about the time of implantation of the embryo into the uterus^{2,3} but it is not known whether this arises by a process of X-chromosome inactivation or by X-chromosome activation⁴. There is now good evidence for the expression of the embryonic genome during preimplantation development⁵⁻¹⁰ and Lyon^{2,3} has reviewed the evidence suggesting that X chromosomes are also expressed during this period. This includes both indirect observations¹¹⁻¹³ and the more direct experiments with the X-linked enzyme hypoxanthine guanine phosphoribosyl-transferase (HPRT) (ref. 10). We have examined the expression of another X-linked enzyme, α -galactosidase¹⁴⁻¹⁵ and report here a 300-fold increase in activity during preimplantation development. This is probably due to the expression of embryonic X chromosomes.

Large numbers of preimplantation mouse embryos at similar stages of development were obtained by artificial insemination (refs 16, 21). Galactosidase activity was assayed in single embryos using the microfluorometric assay technique described for β -glucuronidase^{5,17}. The α -galactosidase assay mixture contained 2×10^{-3} M 4-methylumbelliferyl- α -D-galactopyranoside (Koch-Light), 0.1 M sodium citrate-citric acid (pH 4.2), 0.9% sodium chloride and 0.1% bovine serum albumin.

Embryos were incubated in microdroplets containing 1-8 nl of reaction mixture at either 37 °C or 25 °C for up to 1 h. To avoid depleting the substrate below the optimum concentration, embryos beyond the eight-cell stage were incubated in the reaction mixture for shorter periods (20 min was the shortest incubation time used). The accuracy of the microassay for α -galactosidase was confirmed using a test-tube fluorometric assay^{18,19} for samples of 20-40 embryos. The embryos were added, in less than 5 μ l of phosphate buffered saline, to a reaction volume of 100 μ l which was then frozen and thawed six times and incubated at 37 °C for up to 6 h. The reaction mixture used was the same as for the microassay and at pH 4.2 the reaction was linear with time for 8 h.

The mean α -galactosidase activity per embryo increased 300-fold (from 10^{-14} to 3×10^{-12} mol product per h per embryo at 37 °C) between the two-cell stage (30 h after artificial insemination) and the blastocyst stage (90 h) for C3H/HeHa embryos (Fig. 1). The rate of increase in α -galactosidase activity per embryo also represents a considerable increase in enzyme activity per cell during the initial cleavage divisions (Table 1). This rapid increase in α -galactosidase activity coincides with the increase in activity of HPRT (ref. 10). Other embryonic genes are also expressed at this time⁵ suggesting that the increased α -galactosidase activity is a result of X-chromosome function during early development.

If both X chromosomes of female embryos are active during early preimplantation development X-chromosome dosage compensation would involve the inactivation of one of two active X chromosomes rather than the activation of

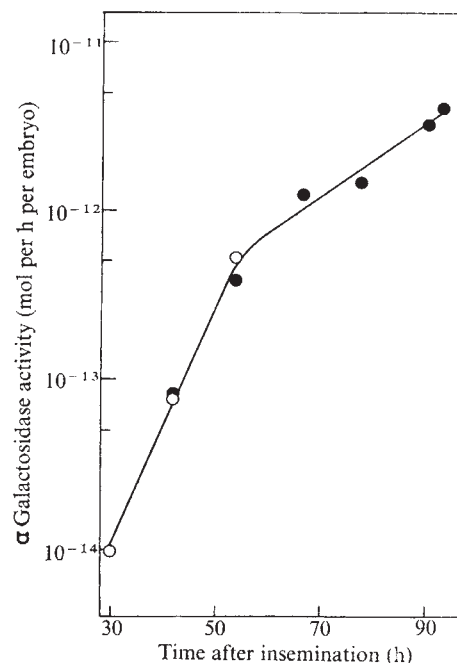


Fig. 1 Increase in mean α -galactosidase activity during preimplantation development of C3H/HeHa mouse embryos. The enzyme activity is expressed as mol of product per h per embryo at 37 °C and results from both microassay (O) and test-tube assay (●) are shown. There is a 300-fold increase in α -galactosidase activity between the 2-cell stage and late blastocyst stage.

one of two inactive chromosomes. It would then be predicted that before X-chromosome inactivation occurs female embryos would have twice the α -galactosidase activity of comparable male embryos, assuming a negligible contribution from the oocyte. The distribution of α -galactosidase activity per embryo for a population of embryos would therefore be expected to be bimodal before X-chromosome inactivation and would become unimodal afterwards.

We have attempted to test this prediction both in order to confirm that both X chromosomes of female embryos are active early in development and to determine whether X-chromosome inactivation occurs before implantation. Bimodal distributions of α -galactosidase activity per embryo were usually obtained for embryos collected at 54 h, and 66 h but in some experiments the distributions were not clearly bimodal. Figure 2 shows the distribution obtained at six time points for our largest experiment in which a possibly bimodal distribution can be seen for the embryos collected at 66 h. Despite our attempts to select embryos at similar stages of development the standard errors for cell counts shown in Table 1 indicate that there is still considerable developmental variation. As the α -galactosidase activity increases very rapidly during this period any

Table 1 Increase in α -galactosidase activity during preimplantation development

Age (h after insemination)	Mean cell no.* ($\pm$ s.e.)	Mean α -galactosidase activity† per cell
30	2	5
42	4	21
54	8	48
66	17.3 ± 0.9	72
78	30.1 ± 2.0	49
90	60.5 ± 4.0	57

*The cell number in 66, 78 and 90 h embryos was determined by the air drying method of Tarkowski²⁰.

†The mean α -galactosidase activity was determined by microassays of 2-cell embryos and test tube assays of older embryos and is expressed as fmol product per cell per h at 37 °C.

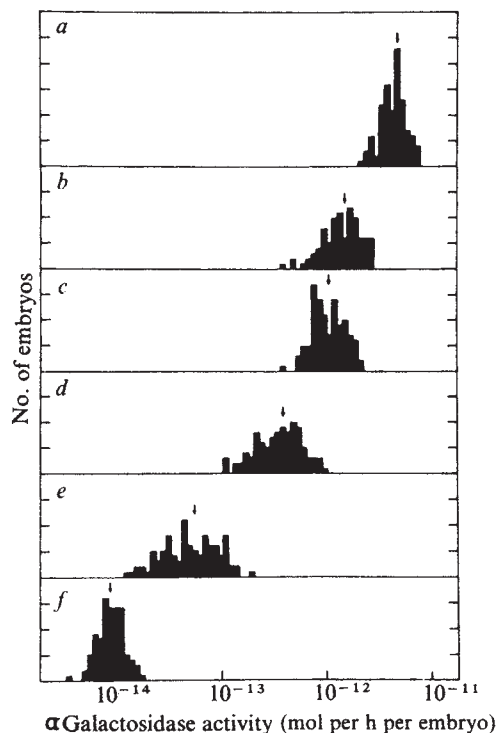


Fig. 2 Distribution of α -galactosidase activity per C3H/HeHa mouse embryo at six time intervals after artificial insemination. Individual embryos were assayed at 25 °C using the microassay and in the experiment illustrated the distribution of α -galactosidase activity per embryo is bimodal for 66-h embryos (morulae). The number of embryos in each activity class is indicated by the vertical scale, marked in intervals of five embryos. The mean enzyme activity for each distribution is indicated with an arrow. a, 90 h, mean α -galactosidase activity ($\pm$ s.e.) $4,523 \pm 124$ fmol product per h per embryo at 25 °C ($n = 104$); b, 78 h, $1,504 \pm 62$ ($n = 91$); c, 66 h, $1,094 \pm 38$ ($n = 111$); d, 54 h, 386 ± 19 ($n = 99$); e, 42 h, 56.1 ± 3.4 ($n = 99$); f, 30 h, 8.7 ± 0.3 ($n = 95$).

developmental variation among embryos of the same age would contribute to variation in enzyme activity per embryo and tend to mask any underlying bimodal distribution.

The striking increase in α -galactosidase strongly suggests that X chromosomes are expressed in preimplantation mouse embryos as early as the four cell stage. We have not yet been able to confirm that female embryos have higher levels of α -galactosidase activity than males during preimplantation development. However, the indications of bimodal distributions of α -galactosidase activity per embryo suggest that both X chromosomes are expressed in female preimplantation embryos by the eight-cell to morula stage. These results support the view that X-chromosome dosage compensation occurs by the inactivation of one of two active X chromosomes rather than by the activation of one of two inactive X chromosomes.

This work was supported by the Lalor Foundation (grant to J.W.) and the US NIH.

DAVID A. ADLER
JOHN D. WEST
VERNE M. CHAPMAN

Department of Molecular Biology,
Roswell Park Memorial Institute,
Buffalo, New York 14263

Received 14 March; accepted 5 May 1977.

- ¹ Lyon, M. F. *Nature* **190**, 372–373 (1961).
- ² Lyon, M. F. *Biol. Rev.* **47**, 1–35 (1972).
- ³ Lyon, M. F. *Proc. R. Soc. B* **187**, 243–268 (1974).
- ⁴ Cattanaach, B. M. *A. Rev. Genet.* **9**, 1–18 (1975).
- ⁵ Wudl, L. & Chapman, V. M. *Dev. Biol.* **48**, 104–109 (1976).
- ⁶ Chapman, V. M., Whitten, W. K. & Ruddle, F. H. *Dev. Biol.* **26**, 153–158 (1971).
- ⁷ Brinster, R. L. *Biochem. Genet.* **9**, 187–191 (1973).

- ⁸ Muggleton-Harris, A. L. & Johnson, M. H. *J. Embryol. exp. Morph.* **35**, 59–72 (1976).
- ⁹ Krco, C. J. & Goldberg, E. H. *Science* **193**, 1134–1135 (1976).
- ¹⁰ Epstein, C. J. *J. Biol. Chem.* **245**, 3289–3294 (1970).
- ¹¹ Epstein, C. J. *Science* **175**, 1467–1468 (1972).
- ¹² Epstein, C. J. *Biol. Reprod.* **12**, 82–105 (1975).
- ¹³ Morris, T. *Genet. Res.* **12**, 125–137 (1968).
- ¹⁴ Burgoyne, P. S. & Biggers, J. D. *Dev. Biol.* **51**, 109–117 (1976).
- ¹⁵ Luthardt, F. W. *Dev. Biol.* **54**, 73–81 (1976).
- ¹⁶ Kozak, C., Nichols, E. & Ruddle, F. H. *Somat. Cell Genet.* **1**, 371–382 (1975).
- ¹⁷ Lusis, A. J. & West, J. D. *Biochem. Genet.* **14**, 849–855 (1976).
- ¹⁸ Leckie, P. A., Watson, J. G. & Chaykin, S. *Biol. Reprod.* **9**, 420–425 (1973).
- ¹⁹ Wudl, L. & Paigen, K. *Science* **184**, 992–994 (1974).
- ²⁰ Kint, J. A. *Science* **167**, 1268–1269 (1970).
- ²¹ Lusis, A. J. & Paigen, K. *Cell* **6**, 371–378 (1975).
- ²² Tarkowski, A. K. *Cytogenetics* **5**, 394–400 (1966).
- ²³ West, J. D. Freis, W. I., Papaioannou, V. E., Karr, J. P. & Chapman, V. M. *J. Embryol. exp. Morph.* (in press).

Low specific activity of rare allozymes of α -glycerophosphate dehydrogenase in *Drosophila*

ELECTROPHORETIC surveys of *Drosophila* species¹ show that many gene-enzyme systems are essentially monomorphic. Very often, however, one or more populations may contain rare alleles, that is, at frequencies less than 5%. Are these alleles selectively neutral and rare because of historical accidents² or are they deleterious and maintained in populations only through mutation? The dimeric enzyme α -glycerophosphate dehydrogenase (α GPDH) is a good example of an essentially monomorphic *Drosophila* gene-enzyme system. Out of the 101 species of *Drosophila* surveyed thus far for electrophoretic variation of α GPDH (refs 3–5), only *D. melanogaster* is polymorphic. This polymorphism seems to be maintained by selection with temperature being, or strongly correlated with, the selective agent^{6–8}. Microcomplement fixation experiments indicate that this enzyme is evolving slowly among *Drosophila* species and has accepted relatively few changes in its structure⁹. We propose that the rare alleles of the α GPDH locus found in some species (see ref. 5) may be at low frequencies because the products of these alleles (allozymes) are catalytically less efficient than their 'wild-type' counterparts.

To test this hypothesis, the specific activities of all the available rare electrophoretic variants (allozymes) were compared with the specific activities of the common allozymes in six species of *Drosophila*. Two other species, *D. nasuta* and *D. birchii*, for which population survey data are not available were also examined. The 'rare' alleles in these species were those found at low frequencies in laboratory stocks. Although there was an insufficient number of flies homozygous for the rare alleles to determine specific activities for the purified allozymes directly, there is a way to determine them indirectly. α GPDH is a dimeric enzyme and in α GPDH^A/ α GPDH^B heterozygotes the A : A, A : B and B : B dimers can be separated electrophoretically. The activities of each of these allozymes can then be measured densitometrically after staining for α GPDH activity. If it is assumed that each allele produces an equal amount of protein and that the monomers associate randomly, then the relative activities of the A : A and B : B homodimers in the heterozygous flies, in fact, represent the relative specific activities of those two forms of α GPDH (Fig. 1).

The critical assumption that the monomers associate randomly is supported by three facts. First, *in vitro* subunit hybridisation experiments using purified α GPDH from eight species of *Drosophila* have shown that the dissociated monomers of these different α GPDHs reassociate randomly¹⁰. Second, theoretical calculations on the effect of single amino acid substitutions in the subunit contact region of this molecule strongly suggest that they would have little discernible effect on the ratio of dimers in heterozygotes¹⁰. Third, only a minority of the many rare variants of haemoglobin affect the subunit contact regions of this molecule¹¹. Thus, few, if any, of the rare allozymes of α GPDH are expected to be affected in the subunit contact

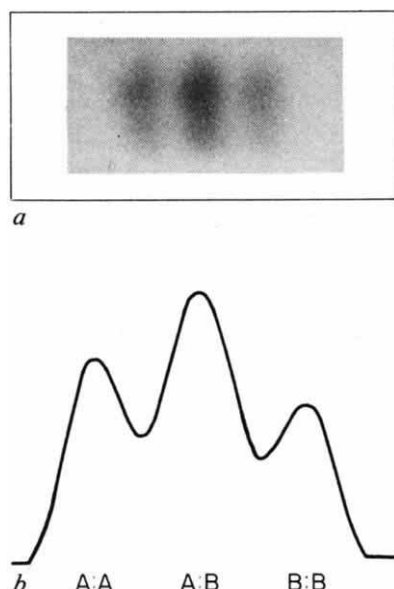


Fig. 1 Electrophoretic separation of α GPDH allozymes from *D. paramelanica* heterozygous at the α GPDH locus. *a*, Cellulose acetate membrane stained for α GPDH activity after electrophoresis. Electrophoresis was carried out in a Beckman Microzone Cell at 300 V for 1 h, in a buffer consisting of one part Tris-barbital-sodium barbital buffer (0.025 ionic strength, pH 8.8) and one part Tris-HCl buffer (0.1 M, pH 8.3). The membranes were stained for 10 min at 20 °C in 25 ml of 0.05 M Tris-HCl buffer, pH 8.3, which contained 15 mg NAD⁺, 5 mg MTT tetrazolium, 3 mg phenazine methosulphate, 200 mg α -glycerophosphate, and 375 mg agar. For the staining time used, the deposition of dye is linear and therefore proportional to enzyme activity (G.E.C., unpublished results). *b*, Densitometer scan of the stained membrane. The area under the B:B peak compared with that under the A:A peak is taken as the relative specific activity of the B:B homodimers.

region of this molecule. It is also doubtful that such effects would alter the ratio of dimers in heterozygotes.

Table 1 lists the observed activity ratios for α GPDH in heterozygotes, together with those expected on the basis of the relative specific activities calculated as outlined. Although there are no gross departures from expectation in these values, there are consistent deficiencies for the

observed values for the heterodimers. The most likely explanation is that the heterodimers do not have specific activities exactly intermediate to those of the respective homodimers, as assumed for calculating the expected ratios in Table 1, but rather have specific activities slightly lower. Thus, in spite of the approximate nature of the calculation, we feel that this is a valid means of estimating the relative specific activities of rare allozymes of α GPDH.

The activities of nine rare allozymes (the B:B allozymes) thus calculated (Table 1) are all lower than their wild-type counterparts (the A:A allozymes). In addition, although the electrophoretic separation of the dimers in α GPDH^A/ α GPDH^B heterozygotes of *D. immigrans* was not sufficient to make a densitometric tracing, it could be seen that the B:B homodimer was not as intensely stained as the A:A homodimer. From these results it might seem that α GPDH^B of *D. birchii* might be a rare allele in natural populations and that α GPDH^B in *D. nasuta* might be polymorphic. Surveys of α GPDH variation in populations of these two species would represent a further test of this hypothesis.

It is also interesting that homozygotes for the rare alleles in some species have low viabilities, reduced fertility and/or longer developmental times. For example, in *D. putrida*, α GPDH^B/ α GPDH^B homozygotes were obtained from matings between heterozygotes, but no adult flies ever eclosed from crosses between these α GPDH^B homozygotes. In *D. tripunctata*, homozygotes for the rare allele were viable but their egg-adult developmental period was 3–4 d longer than wild-type flies raised in the same conditions. The same was true in *D. algonquin*, except that the developmental time of the α GPDH^B homozygotes was only increased 2–3 d. Although the possibility that these effects are due to other closely linked deleterious genes cannot be completely ruled out, it is unlikely that this is the case because none of the α GPDH^A/ α GPDH^A stocks established from single pair matings ever exhibited these characteristics.

The total activity of α GPDH measured on cellulose acetate membranes is virtually the same for α GPDH^A/ α GPDH^A and α GPDH^B/ α GPDH^B homozygotes, although the two homodimers in a heterozygote do not have the same relative activities. This is true even in those species

Table 1 Ratio of α GPDH dimers in heterozygotes and relative specific activities of α GPDH produced by rare alleles

Species	Allele*		Ratio of dimers (% total activity)			Relative specific activity† (B:B against A:A homodimers) (%)	(Range)
			A:A	A:B	B:B		
<i>D. nasuta</i> ¶	B	Observed (s.e.)	28.3 (2.3)	44.7 (2.9)	27.0 (1.8)	96	(90–104)
		Expected‡	25.5	50.0	24.5		
<i>D. robusta</i>	B		29.7 (2.5)	43.9 (4.8)	26.4 (2.6)	89	(79–100)
			26.5	50.0	23.5		
<i>D. robusta</i>	C		32.1 (5.5)	39.3 (10.4)	28.6 (5.5)	89	(81–95)
			26.5	50.0	23.5		
<i>D. affinis</i>	B		29.0 (3.0)	49.3 (3.3)	21.7 (0.8)	75	(66–81)
			28.6	50.0	21.4		
<i>D. paramelanica</i>	B		31.3 (2.2)	45.6 (3.1)	23.1 (1.1)	74	(64–79)
			28.7	50.0	21.3		
<i>D. birchii</i> ¶	B		33.1 (2.1)	43.8 (3.2)	23.1 (1.5)	70	(55–80)
			29.4	50.0	20.6		
<i>D. algonquin</i>	B		36.0 (3.1)	41.0 (4.8)	23.0 (2.2)	64	(58–66)
			30.5	50.0	19.5		
<i>D. tripunctata</i>	B		36.5 (4.0)	43.1 (5.8)	20.4 (1.9)	56	(54–60)
			32.0	50.0	18.0		
<i>D. putrida</i>	B		41.7 (2.1)	47.7 (3.1)	10.6 (1.9)	20	(15–32)
			41.7	50.0	8.3		

*The common allele in each species is designated A. Except for ¶, the remaining alleles occurred in natural populations at frequencies < 0.02 (refs 6,7).

¶The alleles for these species were found in laboratory stocks. Their frequency in wild populations is not known.

‡The expected ratio was calculated using the relative specific activities of the A:A homodimer as $p = 1$, and that of B:B as $q < 1$, (see †). Then the expected ratio is A:A = $p/2(p+q)$; A:B = $(p+q)/2(p+q)$; and B:B = $q/2(p+q)$.

†The method for determining relative specific activities is discussed in the text and Fig. 1. Each recorded value is an average of 4–8 separate determinations.

having a rare allozyme of rather low relative specific activity, for example *D. putrida*. Berger¹² pointed out that the activity level of an enzyme may be an important factor in the maintenance of protein polymorphism. Thus, an individual which has to make more of a less active allozyme may suffer from the biochemical cost of increased synthesis. In this regard, α GPDH is surely a major protein component of the adult fly, for it is homogeneous after only some 500-fold purification of crude extracts of *D. melanogaster*¹³.

Although additional biochemical characterisation of these allelic products is needed, the preliminary results indicate that the enzymes made by the rare alleles at the α GPDH locus are molecules with lowered enzymatic efficiency (although tightly linked *cis* regulatory sites can be associated with different XDH allozymes in *D. melanogaster*)¹⁴. It therefore seems reasonable that these alleles are rare because they are selected against in natural populations. Whether this holds for rare alleles in other *Drosophila* gene-enzyme systems remains to be determined.

GLEN E. COLLIER*

ROSS J. MACINTYRE

Section of Genetics, Development and Physiology,
Plant Science Building,
Cornell University,
Ithaca, New York 14853

Received 28 January; accepted 9 May 1977.

*Present address: Department of Biochemistry, University of Texas Health Science Center, San Antonio, Texas 78284.

- 1 Powell, J. F. *Evol. Biol.* 8, 79-199 (1975).
- 2 Robertson, A. *Genetics* 81, 775-785 (1975).
- 3 Zouros, E. *Nature* 254, 446-448 (1975).
- 4 Barker, J. S. F. & Mulley, J. C. *Evolution* 30, 213-233 (1976).
- 5 Collier, G. E. *Isozyme Bull.* 10, (in the press).
- 6 Berger, E. *Genetics* 67, 121-136 (1971).
- 7 Johnson, F. M. & Schaffer, H. E. *Biochem. Genet.* 10, 149-163 (1973).
- 8 Miller, S., Pearcey, R. W. & Berger, E. *Biochem. Genet.* 13, 175-188 (1975).
- 9 Collier, G. E. & MacIntyre, R. J. *Proc. natn. Acad. Sci. U.S.A.* 77, 684-688 (1977).
- 10 Collier, G. E. thesis, Cornell Univ. (1976).
- 11 Stamatoyanopoulos, G. A. *Rev. Genet.* 6, 47-70 (1972).
- 12 Berger, E. *Am. Nat.* 110, 823-839 (1976).
- 13 Collier, G. E., Sullivan, D. T. & MacIntyre, R. J. *Biochim. biophys. Acta* 429, 316-323 (1976).
- 14 Chovnick, A., Gelbart, W., McCarron, M., Osmond, B., Candido, E. P. M. & Baillie, D. L. *Genetics* 84, 233-255 (1976).

Phenotypic characterisation of a unique non-T, non-B acute lymphoblastic leukaemia cell line

MANY human lymphoblastoid cell lines have been established from subjects with different haemopoietic disorders as well as from normal individuals¹⁻³. In most cases the inability to identify the precise origin of cells which have been established has limited the usefulness of these lines, especially those derived from leukaemic donors. Although there are a few important exceptions to this generalisation (2, 3, 4 and see below) it is significant that no lines have been established from the common form of acute lymphoblastic leukaemia (ALL) which is usually referred to as 'null' cell or 'non-T, non-B' ALL (refs 5-7). We describe here the establishment of a lymphoid cell line, Reh, derived from an acute lymphocytic leukaemia where the leukaemic identity of established cells is supported by kinetic data, chromosome markers and detection of leukaemia associated antigens and enzymes.

Reh cells do not contain detectable Epstein-Barr virus (EBV), either by the nuclear antigen (EBNA) test or by nucleic acid hybridisation (to be published elsewhere). This line was initiated from the peripheral blood of a 15-yr-old girl in relapse with 100% lymphoblasts in her bone marrow and 146,000 white blood cells (mm³) with more than 95% lymphoblasts in her peripheral blood. Clinical and haematological features including morphology and histochemistry of the leukaemic blasts were characteristic of ALL.

Details of culture and karyotype techniques have been described previously⁸. Karyotypes were obtained and studied at the time of seeding the culture, and after establishment without addition of PHA. Fresh leukocytes consisted of populations which were cytogenetically different, one normal: 46, XX, the other abnormal: 45, XX, -2B, +1C (Fig. 1a).

Denaturation techniques have suggested that the two missing B are possibly one number 4 and one number 5 and further indicated that the extra C is not a true C, but probably reflects a structural rearrangement. Early in the course of establishment, chromosome examination revealed that the selection of the abnormal population had

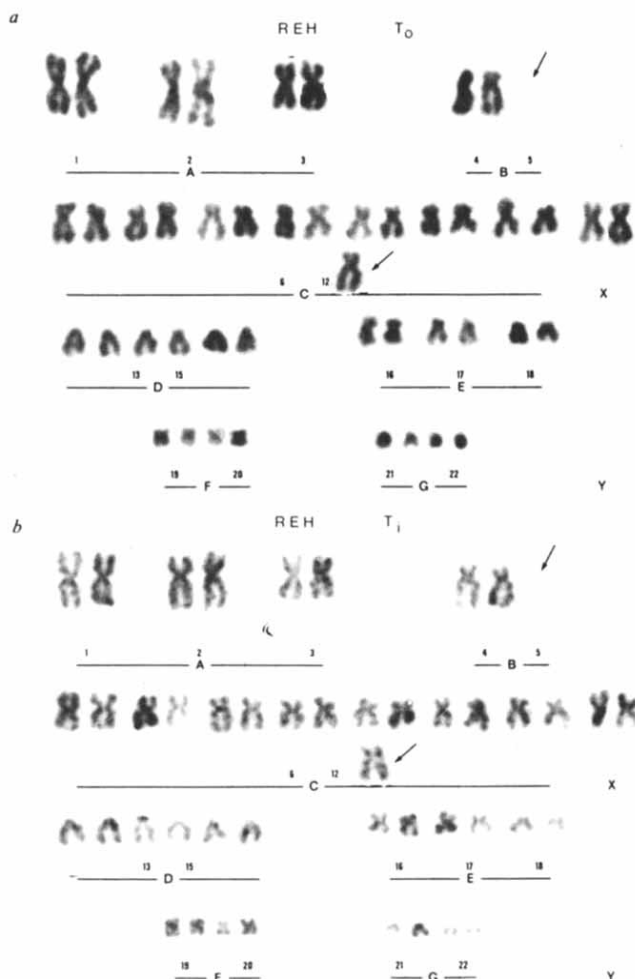


Fig. 1 Chromosome preparations from Reh cells. To, Fresh cells. Ti, Established line. See text for interpretation.

occurred since all the cells showed the 45, XX, -2B, +1C karyotype (Fig. 1b). Disappearance of the normal population was confirmed by further sequential examinations. Analysis with banding techniques is necessary to establish the precise nature of Reh chromosomal characteristics. The results so far, however, indicate clearly that the cells established *in vitro* belong to the same cytogenetically abnormal clone as the patients' uncultured leukaemic cells.

In contrast to the situation found with normal B-cell lines, the Reh cells started to divide at the beginning of the culture. This can be seen in Table 1 from the percentage of mitosis and from the cell uptake of tritiated thymidine (³HTdr). These observations suggest that abnormal leukocytes were immediately able to proliferate in the culture medium and overcome the cytogenetically normal population.

The fresh Reh cells and the Reh line when established were examined for reactivity with a small panel of membrane markers. The initial cells from the patient and the

Table 1 Tritiated thymidine uptake and percentage of mitosis during the establishment of the Reh cell line.

Time in culture (d)	% ³ HtDR labelled cells	% Mitosis
0	14.1	0.13
7	11.4	0.12
14	10.6	0.23
26	9.4	0.17
42	12.1	0.32
54	20.9	0.28
67	13.2	0.45
77	17.5	0.85
96	19.2	0.90
201	22.9	1.10

The percentage of DNA synthesis was determined by the 'dipping' autoradiographic technique. After incubation for 1 h at 37 °C with ³HtDR (1 µCi ml⁻¹ TMM 79 B, CEA, France) the cells in suspension were centrifuged (5 min at 800 r.p.m.), spread and fixed 10 min in methanol. The slides were dipped in Ilford K2 emulsion at 45 °C, dried and set in the dark at 4 °C. Exposure time was 7 d. After development (Kodak D19B, AL4) the slides were rinsed under water, dehydrated for 45 min with ethanol 70 °C and coloured with Giemsa dye. The labelling index was determined by counting 1,000 cells. The percentage of mitosis was measured on 10,000 Giemsa-stained cells before establishment and on 2,000 cells after establishment.

established cell *in vitro* were negative for sheep erythrocyte rosette formation and surface immunoglobulin and could therefore be considered to be of the common type of ALL (that is 'null' ALL or 'non-T, non-B' ALL). Soon after establishment a small proportion of the Reh cells seemed to have cell surface C3 binding sites as determined by the EAC rosette assay⁹; however these were no longer detectable after a few months of culture.

The Reh line has now been maintained in culture for 2 yr and during this time has been used for active immunotherapy. Recently several new leukaemic cell markers have become available^{6,10,11} and so a more extensive phenotyping

of the Reh line has now been carried out. Table 2 shows the results of this analysis in which the phenotype of Reh is compared with that observed in a large series (400) of ALL cases^{6,12}. The phenotype profile of Reh is indeed essentially identical to that of the common form of ALL. The crucial markers here are the anti-ALL sera¹³, anti-p28,33 serum^{12,14} and the enzymes hexosaminidase (intermediate peak isoenzyme)¹² and terminal deoxynucleotidyl transferase¹⁵ since these together constitute a phenotype unique to non-T, non-B ALL¹². Figure 2 illustrates the reactivity of Reh cells with anti-ALL. Absorption of anti-ALL serum with Reh cells removes all binding activity against fresh ALL cells (M.G., unpublished) indicating complete cross-reactivity of the ALL antigen on Reh and non-T, non-B ALL.

Several cell lines have been established from ALL by other investigators. With rare exceptions, however, these have either been EBV positive B cells (presumably derived from the patient's normal B cells) or in cases of Thy or T-like ALL, they have been EBV negative T-cell lines^{16,17}. The latter are presumably leukaemic cell derivatives and in some cases this has been confirmed by chromosome markers. An exception to this generalisation is the establishment of the line NALM-1 from a child with Philadelphia positive CML in blast crisis⁴. The membrane phenotype of this line is the same (with respect to those markers tested) as Reh including the expression of the ALL antigen⁴. This observation accords with recent studies indicating that some cases of CML and non-T, non-B ALL may originate from a common target cell¹⁸. In general it has proved impossible to establish non-T, non-B ALL cell lines (J. Minowada, personal communication). The only reported exceptions in addition to Reh and the Ph¹ positive NALM-1 are two lines of non-T, non-B ALL established from the cells of patients in relapse¹⁹. It may be that the only non-T, non-B ALL type cells which have a capacity

Table 2 Membrane marker and enzyme assays on the Reh line compared with non-T, non-B ALL cells

Markers	Reh line	Non-T, non-B ALL	Ref.*
Antisera			
R. anti-ALL, R6†	+	+	13, 24, 25
R. anti-ALL R73	+	+	
R. anti-ALL R74	+	+	
R. anti-ALL R8 F(ab') ²	+	+	
R. anti-p28, 33 'Ia-like antigen‡	+	+	14, 25
R. anti-monkey thymus F(ab') ² sheep anti-human Ig§	—	—	
Sheep anti-kappa light chain	—	—	25, 26
Sheep anti-lambda light chain	—	—	
R. anti-'myeloid' antigen R. 101	—	—	27, 28
antigen R. 102	—	—	
Rosette assays			
E sheep	—	—	29, 30
E mouse: no neuraminidase	—	—	
E mouse: plus neuraminidase	+(39%)	—(?)††	
E (Ox) IgM	—	—	31
E (Ox) IgM + C	—	—	31
E (Ox) IgG	—	—	31
Enzymes			
Hexosaminidase 'I' isoenzyme	++	+/++++	12, 32
Terminal deoxynucleotidyl transferase	360**	1.5–190 (mean 57)	15, 33

*Reference for reagents¹, assay methods, and results on uncultured ALL cells.

†R6, 73, 74, 8 are for different rabbit anti-ALL sera. They are extensively absorbed with normal tissues and non-ALL leukaemias.

‡A gift from Dr Ken Welsh (Queen Victoria Hospital, East Grinstead).

§A gift from Dr I. Chantler (Burroughs-Wellcome).

||Very faint staining with both anti-kappa and anti-lambda was observed on all Reh cells. We interpret this as minimal binding which is probably nonspecific.

¶Rabbit anti-acute myelo-monocytic leukaemia cell serum absorbed with lymphoid cells. This serum defines a myeloid differentiation antigen which is absent from normal and leukaemic lymphoid cells (Roberts, M. and M.F.G., in preparation).

**Units per 10⁶ cells. 1 unit is 1 nmol⁻¹ of nucleotide polymerised at 37 °C (Ganeshaguru, K., Hoffbrand, G. V., Janossy, G. and M.F.G., in preparation). E, erythrocyte. C, complement. Fresh mouse serum used for this complement receptor assay.

††Out of 6 tested.

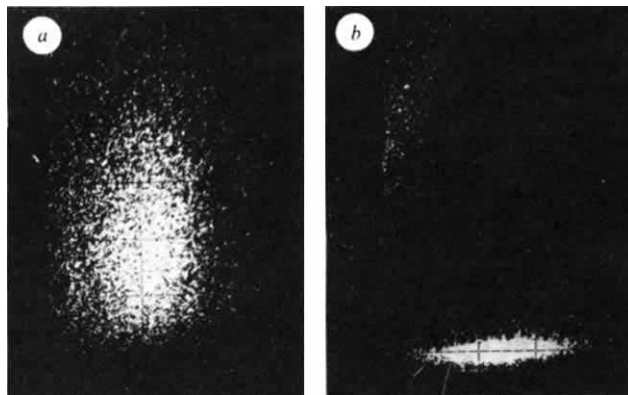


Fig. 2 Fluorescence activated cell sorter (FACS) analysis of anti-ALL binding by Reh line. Vertical axis: relative fluorescence intensity. Horizontal axis: relative cell size. *a*, Reh cells stained with rabbit anti-ALL followed by fluorescein labelled goat anti-rabbit Ig. *b*, Reh cells treated with normal rabbit serum followed by fluorescein labelled goat anti-rabbit Ig. See ref. 6 for details of FACS and fluorescence staining methodology.

for unlimited growth *in vitro* are those which are aneuploid with major chromosomal alterations. These changes may occur at the onset of the disease and affect all members of the leukaemic clone or more frequently may arise after treatment in a subclone (perhaps in Ph⁺ negative ALL). The ability of these exceptional ALL or CML blast crisis cells to grow *in vitro* may be a correlate of their growth patterns *in vivo* since relapsed ALL is often more drug resistant than previously untreated ALL (ref. 20) and Ph⁺ positive disease, even when presenting with an ALL phenotype, has a worse prognosis than Ph⁺ negative ALL in the same age group^{18,21,22}.

The availability of cell lines which are clearly derived from the common form of lymphoblastic leukaemia may be of considerable value in further investigations on the nature of leukaemia-associated antigens, and the block to further differentiation which presumably exists in these cells. The Reh line has been used extensively for active immunotherapy of ALL patients in Villejuif²³. It is possible that the therapeutic efficacy of Reh might be different from that of fresh allogeneic ALL cells. The ALL associated antigen might in combination with allogeneic histocompatibility antigens provide a suitable target for the patients' immune system. We are currently testing sera from treated patients to see if antibodies are present against the ALL antigen which the rabbit antisera recognise.

This work was supported by ATP INSERM and the Imperial Cancer Research Fund. We thank Dr R. Ellis (Institute for Child Health) for hexosaminidase assays and Professor G. V. Hoffbrand and Dr K. Ganeshaguru (Royal Free Hospital) for terminal transferase enzyme assays.

C. ROSENFELD
A. GOUTNER
C. CHOQUET
A. M. VENUAT
B. KAYIBANDA
J. L. PICO

*Institut de Cancerologie et d'Immunogenetique,
INSERM et Association Claude Bernard,
16, Avenue Paul Vaillant Couturier,
Villejuif, France*

*Membrane Immunology Laboratory,
Imperial Cancer Research Fund,
Lincoln's Inn Fields,
London, UK*

M. F. GREAVES

Received 11 February; accepted 19 April 1977.

- 1 Moore, G. E. & Minowada, J. in *Hemic Cells in Vitro* (ed. Farnes, P.) 110 (The Williams and Wilkins, Baltimore, 1969).
- 2 Nilsson, K. & Ponten, J. *Int. J. Cancer* **15**, 321 (1975).
- 3 Zech, L., Haglund, U., Nilsson, K. & Klein, G. *Int. J. Cancer* **17**, 47 (1976).
- 4 Minowada, J., Tsubota, T., Greaves, M. F. & Walters, T. R. *J. natn. Cancer Inst.* (in the press).
- 5 Brouet, J.-C. *et al. Br. J. Haematol.* **33**, 319 (1976).
- 6 Greaves, M. F. *Progr. Haematol.* **9**, 255 (1975).
- 7 Collins, R. D., Smith, J. L., Klein, G. P. *Br. J. Haematol.* **26**, 615 (1974).
- 8 Rosenfeld, C., Macieira-Coelho, A., Venuat, A. M., Jasmin, C. & Tuan, T. Q. *J. natn. Cancer Inst.* **43**, 581 (1969).
- 9 Ehlenberger, A. G. & Nussenzweig, V. in *In Vitro Methods of Cell Mediated and Tumor Immunity* (eds Bloom, B. & David, J.) 113 (Academic, New York, 1976).
- 10 Thierfelder, S., Rodt, H. & Thiel, E. (eds) *Immunological Diagnosis of Leukaemias and Lymphomas* (Lehmanns, Munich and New York).
- 11 Special WHO-IUIS technical report *Br. J. Haematol.* (in the press).
- 12 Greaves, M. F. in *Immunological Diagnosis of Leukaemias and Lymphomas* (eds Thierfelder, S., Rodt, H. & Thiel, E.) 61 (Lehmanns, Munich and New York, 1977).
- 13 Brown, G., Capellaro, D. & Greaves, M. F. *J. Natn. Cancer Inst.* **55**, 1281 (1975).
- 14 Schlossman, S., Chess, L., Humphreys, R. E. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1288 (1976).
- 15 McCaffrey, R., Harrison, T. A., Parkman, R. & Baltimore, D. *New Engl. J. Med.* **292**, 775 (1975).
- 16 Kaplan, J., Shope, T. C. & Peterson, W. D. *J. exp. Med.* **139**, 1070 (1974).
- 17 Minowada, J. *et al.* in *Immunological Diagnosis of Leukaemias and Lymphomas* (eds Thierfelder, S., Rodt, H. & Thiel, E.) (Lehmanns, Munich and New York, in the press).
- 18 Janossy, G., Roberts, M. & Greaves, M. F. *Lancet* **ii**, 1058 (1976).
- 19 Schneider, U., Schwenk, H. U. & Bornkamm, G. in *Immunological Diagnosis of Leukaemias and Lymphomas* 265 (eds Thierfelder, S., Rodt, H. & Thiel, E.) (Lehmanns, Munich and New York, 1977).
- 20 Gunz, F. & Baiki, A. G. (eds) *Leukaemia* (Grune and Stratton, New York, 1974).
- 21 Secker-Walker, L. M. *et al. Lancet* **ii**, 1405 (1976).
- 22 Greaves, M. F. in *Immunodiagnosis of Cancer* (ed. Herbermann, R.) (Academic, New York, in the press).
- 23 Mathé, G. *et al. Proc. R. Soc. Med.* **68**, 211 (1975).
- 24 Brown, G., Hogg, N. & Greaves, M. F. *Nature* **258**, 454 (1975).
- 25 Janossy, G., Greaves, M. F., Sutherland, R., Durrant, J. & Lewis, C. *Leukemia Research* **2**, (in the press).
- 26 Greaves, M. F. & Janossy, G. in *In Vitro Methods in Cell Mediated and Tumor Immunity* (eds Bloom, B. & David, J.) (Academic, New York, 1976).
- 27 Greaves, M. F., Falk, J. & Falk, R. *Scand. J. Immun.* **4**, 80 (1975).
- 28 Baker, R., Falk, J., Falk, R. & Greaves, M. F. *Br. J. Haematol.* **32**, 13 (1976).
- 29 Catovsky, D., Cherchi, M., Okos, A., Hegde, O. & Galton, D. A. G. *Br. J. Haematol.* **33**, 173 (1976).
- 30 Stathopoulos, G. & Davies, A. J. S. *Lancet* **i**, 1078 (1976).
- 31 Bloom, B. & David, U. (eds) *In Vitro Methods of Cell Mediated and Tumor Immunity* (Academic, New York, 1976).
- 32 Ellis, R. B., Ikonne, J. U. & Masson, P. K. *Analyt. Biochem.* **63**, 5 (1975).
- 33 Hutton, J. J. & Coleman, M. S. *Br. J. Haematol.* **37**, 447 (1976).

Human B cell, T cell and null cell leukaemic cell lines derived from acute lymphoblastic leukaemias

CONTINUOUS culture of human leukaemic cells is still a difficult task, although several leukaemic cell lines have been derived from patients with acute and chronic myelogenous leukaemia^{1,2} and acute lymphoblastic leukaemia (ALL) (refs 3–6). We describe here B-cell, T-cell and null-cell leukaemic cell lines newly established from three patients with ALL.

Leukaemic cells separated by dextran sedimentation from peripheral blood or bone marrow were resuspended in growth medium at a concentration of 1×10^7 cells per ml and dispersed into 35-mm Petri dishes (2.5 ml per dish). The cultures were incubated at 37 °C in a humidified atmosphere of 7.5% CO₂ in air. Two kinds of growth media were used: medium I consisted of RPMI 1640 supplemented with 20% human cord serum (HCS) and medium II of RPMI 1640 supplemented with 20% foetal calf serum (FCS). Partial medium changes were done twice a week. After cell line establishment, subcultures were made usually once a week.

Cells were examined for receptors for sheep erythrocytes (E), Fc fragment of IgG (EA) and complement (EAC). They were also tested for surface immunoglobulin (Ig) by direct immunofluorescence with FITC-conjugated goat antisera (Behringwerke) monospecific for human IgM, IgG, IgA and light chains. All these assays were performed as described by Jondal and Klein⁷. Epstein-Barr virus (EBV)-determined nuclear antigens (EBNA) was tested for by anti-complement immunofluorescence⁸. Karyotypes were analysed as described previously⁹.

A B-cell line (BALL-1) was derived from peripheral blood of a 75-yr-old man who died 1.5 months after the onset of ALL. Culture was initiated on 15 January 1976

and established only in medium I 1 month later. The peripheral leukocyte count at the beginning of culture was 25×10^4 per cm^3 with 83% lymphoblasts. Both peripheral lymphoblasts and BALL-1 cells were virtually 100% surface Ig-positive for two heavy chain classes IgM and IgA. The light chain was of the lambda type. Cells from both sources contained 36–60% EA and EAC rosetting cells but neither of them formed E rosettes. Autopsy on this patient revealed no lymphadenopathy or tumorous masses.

A T-cell line (TALL-1) was established from bone marrow of a 28-yr-old man who developed the terminal leukaemic phase of lymphosarcoma. Marrow cells cultured on 13 April 1976 grew equally well in both medium I and II after 1 month. The marrow at the time of culture contained 94% lymphoblasts. The majority of marrow lymphoblasts and TALL-1 cells formed E rosettes but neither of them bore surface Ig.

A null-cell line (NALL-1) was obtained from peripheral blood of a 14-yr-old boy who relapsed after a 3-month remission. Culture was started on 16 March 1976 and established only in medium I 2 weeks later. At the time of culture initiation, the peripheral leukocyte count was 18.8×10^4 per cm^3 with 98% lymphoblasts. Neither peripheral lymphoblasts nor NALL-1 cells demonstrated E rosette formation or surface Ig. Receptors for EAC but not for EA were detected on a small percentage (less than 8.1%) of NALL-1 cells, however.

Cells from BALL-1, TALL-1 and NALL-1 grew in single cell suspension with a doubling time of 60–90 h and were morphologically identical to the original lymphoblasts of the respective donor. These cell lines were all repeatedly negative for EBNA when tested by us and in the laboratory of Dr Yorio Hinuma, and hence considered EBV genome-free.

Karyotypic analysis was performed on these cell lines between 7 and 9 months after culture initiation. Fifty metaphases were studied from each sample and several selected metaphases were further analysed by the Giemsa banding technique. BALL-1 had a pseudodiploid mode with four characteristic markers (one each of large telocentric and metacentric chromosomes and two medium-sized telocentric chromosomes). Two metaphases were obtained from peripheral lymphoblasts from the BALL-1 donor, both of which were pseudodiploid. The chromosomes of the TALL-1 line were all in the hypertetraploid region, ranging from 95 to 101. No chromosome preparations were made directly from the donor of the TALL-1 line. NALL-1 had a hypodiploid mode of 43 with a minute marker and missing sex chromosomes. Peripheral lymphoblasts from the NALL-1 donor yielded three analysable metaphases, which all showed 43 chromosomes and the minute marker observed in the NALL-1 line.

When BALL-1, TALL-1 and NALL-1 were stained with rabbit anti-human thymocyte serum by indirect membrane immunofluorescence, only TALL-1 cells were fluorescent, indicating the absence of thymus-related antigen on BALL-1 and NALL-1 cells. The anti-thymocyte serum had been extensively absorbed with 'normal' lymphoblastoid cell lines, chronic lymphocytic leukaemia cells and normal peripheral leukocytes, and known to react specifically with MOLT-4 (leukaemic T-cell line) and thymocytes. Furthermore, BALL-1 and NALL-1 were found to stimulate allogeneic lymphocytes in one-way mixed lymphocyte reactions whereas TALL-1 was non-stimulatory.

Continuous multiplication of B-cell and null-cell leukaemic cells in medium I in contrast to their rapid degeneration in medium II suggests that HCS is more growth-supporting than FCS for certain human ALL cells. On the other hand, T-cell lymphoblasts grew equally well in both medium I and II; this is consistent with the establishment of T-cell ALL lines in other laboratories in

medium containing FCS (refs 3–6). Both BALL-1 and NALL-1 initially established in medium I were then adapted to grow in medium II, although they tended to multiply better in medium I than in medium II. Another advantageous point of the supplemental use of HCS is that antibody present in HCS against EBV capsid antigen (VCA) would suppress transformation of coexisting normal B lymphocytes by EBV (ref. 10). The pooled HCS that we used in the present experiment had an anti-VCA titre of 1:160.

The assumption that all of BALL-1, TALL-1 and NALL-1 originated from leukaemic cells is based on the findings that they uniformly lack the EBV genome, have morphology and surface characteristics identical to those of the leukaemic cells of the corresponding donor and have various chromosome abnormalities including those found in the pre-culture lymphoblasts. Therefore, these cell lines are distinct from EBV-carrying lymphoblastoid cell lines of normal B-lymphocyte origin¹¹.

A question may still be raised as to the possible derivation of the BALL-1 line from normal B lymphocytes of the leukaemic patient, even though it is EBV-negative. This would be unlikely, however: rare cases of B-cell ALL do exist^{12,13}; EBV-negative lymphoblastoid cell lines have never been established from normal tissues¹¹; and in addition to the characteristics described above; BALL-1 cells have produced serially transplantable lethal tumours in newborn hamsters treated with anti-lymphocyte serum (unpublished data).

Minowada *et al.*¹⁴ have established a Ph¹ chromosome-positive non-T non-B (null-cell) line from a child with chronic myelogenous leukaemia in blast crisis. Clearly, this cell line is different from NALL-1 with respect to the Ph¹ chromosome and its derivation. The B-cell and null-cell leukaemic cell lines we describe here are unique and, used together with T-cell lines, should contribute to further understanding of human lymphocyte subpopulations in normal and leukaemic states.

This work was supported by the Ministry of Health and Welfare and the Ministry of Education of Japan. We thank Drs Takeshi Sairenji and Yorio Hinuma for performing the EBNA test.

ISAO MIYOSHI*
SHUNKICHI HIRAKI
TERUHIKO TSUBOTA
ICHIRO KUBONISHI
YUZO MATSUDA
TAKASHI NAKAYAMA
HIROFUMI KISHIMOTO
IKURO KIMURA

Department of Medicine,

HIROSHI MASUJI

Cancer Institute Division of Pathology,
Okayama University Medical School,
Okayama 700, Japan

Received 30 December 1976; accepted 5 May 1977.

* To whom correspondence should be addressed.

- 1 Lozzio, C. B. & Lozzio, B. B. *Blood* **45**, 321 (1975).
- 2 Gallagher, R. E., Salahuddin, S. Z., Hall, W. T., McCredie, K. B. & Gallo, R. C. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4137 (1975).
- 3 Foley, G. E. *et al. Cancer* **18**, 522 (1965).
- 4 Minowada, J., Ohnuma, T. & Moore, G. E. *J. natn. Cancer Inst.* **49**, 891 (1972).
- 5 Kaplan, J., Shope, T. C. & Peterson, W. D., Jr *J. exp. Med.* **139**, 1070 (1974).
- 6 Minowada, J. & Moore, G. E. in *Comparative Leukemia Research. Leukemogenesis* (eds Ito, Y. & Dutcher, R. M.) 250 (University of Tokyo Press, Tokyo, 1975).
- 7 Jondal, M. & Klein, G. *J. exp. Med.* **138**, 1365 (1973).
- 8 Reedman, B. M. & Klein, G. *Int. J. Cancer* **11**, 499 (1973).
- 9 Miyoshi, I. *et al. Int. J. Cancer* **18**, 67 (1976).
- 10 Rickinson, A. B., Jarvis, J. E., Crawford, D. H. & Epstein, M. A. *Int. J. Cancer* **14**, 704 (1974).
- 11 Klein, G. *New Engl. J. Med.* **293**, 1353 (1975).
- 12 Davey, F. R. & Gottlieb, A. J. *Am. J. clin. Pathol.* **62**, 818 (1974).
- 13 Brouet, J. C. *et al. Br. J. Haemat.* **33**, 319 (1976).
- 14 Minowada, J., Tsubota, T., Greaves, M. F. & Walters, T. R. *J. natn. Cancer Inst.* (in the press).

In vitro immunisation against human tumour cells with bacterial extracts

LYMPHOCYTES from normal persons can become immunised to a variety of tumour cells by exposure to tumour cells *in vitro*¹⁻³. Since close antigenic relationships between micro-organisms and some tumour cells have been demonstrated⁶⁻⁹, studies were carried out to see if bacterial antigens could similarly immunise lymphocytes from normal persons *in vitro* against some tumour cells. The experiments reported here indicate that not only is this possible, but that the cytotoxicity of bacterial immunised lymphocytes is sometimes greater than when lymphocytes are sensitised by intact tumour cells.

Soluble bacterial extracts were prepared from the Glaxo strain of *Mycobacterium bovis* (BCG), *Listeria monocytogenes* (LM), *Escherichia coli* (EC), and *Staphylococcus aureus* (SA), as previously described¹⁰ and are referred to as BCG-SS, LM-SS, EC-SS, and SA-SS, respectively. Their nitrogen (N) concentrations were determined¹¹ and they were passed through a Millipore filter before use.

Breast tumour cells that originated from the pleural effusion of a patient with breast carcinoma were obtained from Dr G. Seman, MD Anderson Hospital, Houston, Texas¹². Malignant melanoma cells (RPMI-7932 and RPMI-8322) were provided by Dr G. E. Moore, Denver General Hospital, and are referred to as melanoma-1 and melanoma-2. Melanoma-1 originated from a pleural effusion and melanoma-2 was established from a metastatic solid tumour lesion. Tumour cells were maintained in tissue culture¹.

Peripheral blood lymphocytes were obtained from 12 normal tumour-free persons. Lymphocytes were then isolated and purified over Ficoll-Hypaque⁴. The purified lymphocytes contained less than 5% polymorphonuclear leukocytes and 5-10% monocytes. These were sensitised with mitomycin C treated tumour cells as previously described⁴ or with various concentrations of the bacterial extracts. The cytotoxicity of *in vitro* sensitised cells was then measured by a ⁵¹Cr release assay¹⁻⁴. Results were compared with the capacity of the same cells to kill normal lymphoblastic target cells. In only one case (donor number 4) was there as much as 25% specific ⁵¹Cr release from normal lymphoblastic target cells. In accordingly, ⁵¹Cr release of 25% or more was chosen in this report as the criterion for specific cytotoxicity.

BCG-sensitised lymphocytes from eleven donors caused a greater release of ⁵¹Cr (28% to 78%) from melanoma-1 cells than from the control lymphoblasts (-2% to 21%). Melanoma-2 cells were killed by BCG-immunised lymphocytes from two of three donors and breast tumour cells were killed by BCG-immunised lymphocytes from five of eleven donors. LM-sensitised cells from all seven donors were similarly cytotoxic to melanoma-1 cells and LM-sensitised cells from five of seven donors were cytotoxic to breast tumour cells. In contrast, lymphocytes from only three of seven donors treated with SA-SS became cytotoxic to melanoma-1, and two of seven became cytotoxic to breast tumour cells. Lymphocytes from four donors were treated with EC-SS and none of these developed appreciable cytotoxicity against either melanoma or breast tumour cells. Generally, LM-SS appeared on a nitrogen basis to be more immunogenic than BCG-SS. In some cases, as little as 2 µg N of LM-SS effectively immunised normal lymphocytes. Data from individual experiments where lymphocytes from normal donors were treated with equivalent amounts of BCG, LM, and EC are presented in Fig. 1.

Lymphocytes from some of the donors that were pre-treated with phytohaemagglutinin (PHA), pokeweed mito-

gen (PWM), and concanavalin A (con A) were also treated as effector cells. PHA treated lymphocytes from four of ten persons became cytotoxic to melanoma-1 cells and cells from three of five donors that were PWM treated became cytotoxic to melanoma-1 and one of these also to breast tumour cells. Lymphocytes from one of six donors treated with con A became cytotoxic to melanoma-1 and none to breast tumour cells. PHA treated lymphocytes have

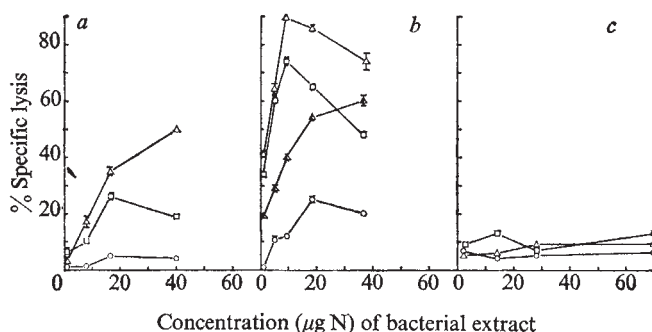


Fig. 1 *In vitro* immunisation of lymphocytes against human tumour with bacterial extracts. Lymphocytes were sensitised with different concentrations of: a, BCG extract (BCG-SS); b, *Listeria monocytogenes* extract (LM-SS), c, *Escherichia coli* (EC-SS). Sensitised lymphocytes were tested against target cells: Δ, melanoma-1 tumour cells; ▲, melanoma-2 tumour cells; □, breast tumour cells; ○, normal lymphoblasts. % Specific lysis of ⁵¹Cr-tumour cells was measured. Only standard deviations greater than 1% are indicated.

previously been shown to be cytotoxic to some tumour cells¹³⁻¹⁶. The degree of cytotoxicity observed in this study was relatively less and more variable with the mitogens tested than those produced by the bacterial extracts, however, especially against breast tumour cells. This would tend to minimise the possibility that the cytotoxicity initiated by the bacterial extracts was entirely due to mitogenic effects.

If *in vitro* sensitisation of lymphocytes by bacterial

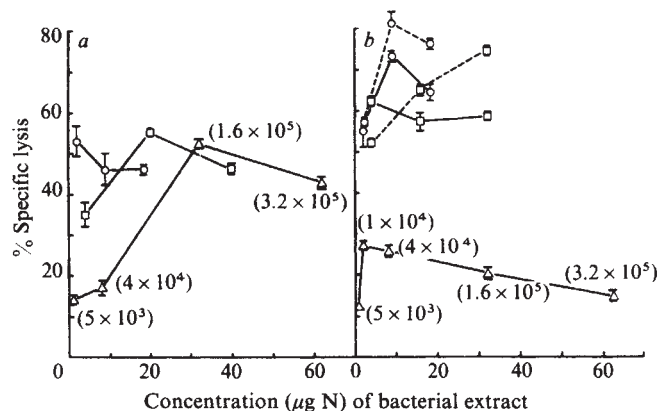


Fig. 2 *In vitro* immunisation of lymphocytes with tumour cells or bacterial extracts. a, Lymphocytes were sensitised with breast tumour cells (Δ), BCG-SS (□), or LM-SS (○) and tested against breast tumour cells. Number of immunising tumour cells shown in parenthesis. b, Lymphocytes sensitised with melanoma-1 cells (Δ), BCG-SS (□), or LM-SS (○) and tested against melanoma-1 cells (Δ). Number of immunising tumour cells in parenthesis. When anti-BCG coated melanoma cells and anti-*Listeria* coated melanoma cells were the target cells, results are indicated by broken lines. % Specific lysis of ⁵¹Cr-tumour cells measured in each case.

extracts is related to antigenic similarities, it is conceivable that three tumour cell lines tested may also have common antigenic determinants. This possibility has been tested and the results suggest this is the case (manuscript submitted for publication).

The possibility exists that the cytotoxic reactions observed between the effector and target cells were due to histocompatibility differences. When lymphocytes were tested

as target cells, however, there was considerably less cytotoxicity and sometimes no specific ^{51}Cr release. Lymphocytes from three donors and the breast tumour cells were kindly typed for HL-A antigens by Dr Robert McCalmon, Veterans Hospital, Denver, Colorado. Melanoma-1 was previously typed in Dr Paul I. Terasaki's laboratory, University of California, Los Angeles. No HL-A antigens were found on breast cells and only HL-A 2 was found

Table 1 *In vitro* immunisation of lymphocytes against human tumour cells with bacterial extracts

Lymphocytes from donor no.	Sensitising tumour cell or bacterial extract*	% Specific ^{51}Cr release $\pm$ s.d. from target cells†			
		Breast	Mel-1	Mel-2	Normal lymphoblasts
1	A $\dagger$	Breast	51 $\pm$ 1.0		
		BCG-SS	24 $\pm$ 0.5		5 $\pm$ 2.0
		LM-SS	25 $\pm$ 1.0		6 $\pm$ 1.0
		SA-SS	39 $\pm$ 1.0		10 $\pm$ 2.0
	A $_1$ †	Breast	44 $\pm$ 2.0		
		BCG-SS	26 $\pm$ 0.5	23 $\pm$ 1.0	-2 $\pm$ 3.0
		LM-SS	31 $\pm$ 2.0	37 $\pm$ 1.0	-15 $\pm$ 3.0
2	Breast	52 $\pm$ 1.0			
		BCG-SS	15 $\pm$ 1.0		7 $\pm$ 1.0
		SA-SS	29 $\pm$ 2.0		0 $\pm$ 1.0
		EC-SS	4 $\pm$ 0.5		2 $\pm$ 0.5
3	Breast	59 $\pm$ 1.0			
		BCG-SS	32 $\pm$ 1.0	50 $\pm$ 1.0	
		LM-SS	48 $\pm$ 2.0	50 $\pm$ 2.0	
4	Breast	33 $\pm$ 1.0			
		Mel-2		33 $\pm$ 1.0	
		BCG-SS	46 $\pm$ 1.0	38 $\pm$ 1.0	21 $\pm$ 0.5
		LM-SS	72 $\pm$ 1.0	60 $\pm$ 1.0	25 $\pm$ 1.0
5	Breast	62 $\pm$ 3.0			
		BCG-SS	34 $\pm$ 1.0	60 $\pm$ 4.0	
		LM-SS	17 $\pm$ 1.0	50 $\pm$ 3.0	
6	Mel-1		24 $\pm$ 4.0		
		SA-SS	11 $\pm$ 1.0	9 $\pm$ 1.0	
7	Breast	52 $\pm$ 1.0			
		Mel-1	27 $\pm$ 0.5		
		BCG-SS	55 $\pm$ 1.0	62 $\pm$ 0.5	7 $\pm$ 1.0
		LM-SS	53 $\pm$ 1.0	73 $\pm$ 1.0	-3 $\pm$ 0.5
8	Breast	28 $\pm$ 1.0			
		BCG-SS	16 $\pm$ 0.5	34 $\pm$ 1.0	-1 $\pm$ 0.5
		SA-SS	5 $\pm$ 1.0	20 $\pm$ 1.0	6 $\pm$ 0.5
		EC-SS	5 $\pm$ 0.5	7 $\pm$ 1.0	-3 $\pm$ 1.0
9	BCG-SS	9 $\pm$ 0.5	56 $\pm$ 3.0		5 $\pm$ 1.0
		SA-SS	7 $\pm$ 0.5	15 $\pm$ 1.0	
		EC-SS	7 $\pm$ 1.0	29 $\pm$ 0.5	-1 $\pm$ 1.0
10	BCG-SS	5 $\pm$ 1.0	47 $\pm$ 1.0		
		LM-SS	9 $\pm$ 0.5	54 $\pm$ 1.0	
		SA-SS	3 $\pm$ 0.5	13 $\pm$ 1.0	
11	BCG-SS	15 $\pm$ 0.5	78 $\pm$ 0.5		11 $\pm$ 1.0
		SA-SS	8 $\pm$ 1.0	11 $\pm$ 1.0	8 $\pm$ 1.0
12	Breast	15 $\pm$ 0.5			
		BCG-SS	19 $\pm$ 1.0	47 $\pm$ 0.5	5 $\pm$ 1.0
		EC-SS	10 $\pm$ 0.5	17 $\pm$ 2.0	0 $\pm$ 1.0

Lymphocytes were sensitised in 0.75 ml of minimal essential medium containing 20% heat inactivated human AB plasma. Fewer than 1% monocytes were present in sensitised cell population. Cytotoxic activity of sensitised cells was tested against breast, melanoma-1 (mel-1), melanoma-2 (mel-2), and normal lymphoblast cells in ^{51}Cr release assay in triplicate. Normal target cells were normal lymphocytes that have been transformed by PHA or PWM.

*The range of bacterial concentrations used was 18 μg to 36 μg N.

†The adjusted percentage of specific ^{51}Cr release was calculated as:

$$\frac{\text{test release} - \text{minimum release}}{\text{maximum release} - \text{minimum release}} \times 100$$

Controls (minimum values) for each experiment included incubation of target cells alone or target cells with normal unsensitised lymphocytes. Control values were 8–16% for breast, 12–22% for melanoma-1, 18–20% for melanoma-2, and 10–16% for PHA-treated normal lymphoblasts. There was no significant difference between background calculated from target cells incubated alone or target cells incubated with lymphocytes which were cultured alone for 5 d. Maximum values (freeze-thaw $\times$ 4) varied between 70 to 83% (average 77%) for melanoma-1, 66 to 72% for melanoma-2, and 68 to 82% (average 72%) for breast tumour cells.

‡A and A $_1$ are two different bleedings from donor 1.

Table 2 Effect of *in vitro* stimulation of lymphocytes with plant mitogens

Lymphocytes from donor no.	Sensitising plant mitogens	% Specific ⁵¹ Cr release $\pm$ s.d. from target cells			
		Breast	Mel-1	Mel-2	Normal lymphoblasts
1	A† PHA	1 $\pm$ 1.0	36 $\pm$ 3.0		-10 $\pm$ 2.0
	A ₁ † PHA	8 $\pm$ 0.5	61 $\pm$ 1.0	15 $\pm$ 1.0	-2 $\pm$ 1.0
	Con A	-2 $\pm$ 0.5	13 $\pm$ 0.5	8 $\pm$ 0.5	-3 $\pm$ 1.0
2	PHA	1 $\pm$ 1.0	15 $\pm$ 1.0		
	PWM	5 $\pm$ 0.5	42 $\pm$ 1.0		
	Con A	1 $\pm$ 0.5	9 $\pm$ 1.0		
3	PHA	13 $\pm$ 1.0	26 $\pm$ 1.0	10 $\pm$ 2.0	
	PWM	51 $\pm$ 1.0	73 $\pm$ 2.0	49 $\pm$ 1.0	
	Con A	3 $\pm$ 0.5	28 $\pm$ 1.0	24 $\pm$ 0.5	
4	PHA	11 $\pm$ 1.0	50 $\pm$ 0.5	20 $\pm$ 0.5	
5	PHA	0 $\pm$ 1.0	9 $\pm$ 0.5		
	PWM	15 $\pm$ 0.5	35 $\pm$ 1.0		
	Con A	6 $\pm$ 1.0	12 $\pm$ 1.0		
6	PHA	1 $\pm$ 0.5	9 $\pm$ 2.0		
	PWM		-11 $\pm$ 1.0		
	Con A		-12 $\pm$ 1.0		
8	PHA	11 $\pm$ 0.5	24 $\pm$ 0.5		
9	PHA	4 $\pm$ 1.0	21 $\pm$ 1.0		
	PWM	8 $\pm$ 1.0	6 $\pm$ 1.0		
	Con A	-1 $\pm$ 0.5	4 $\pm$ 0.5		
10	PHA	3 $\pm$ 1.0	20 $\pm$ 1.0		0 $\pm$ 1.0

Lymphocytes were sensitised with PHA, PWM, or con A for 5 d and cytotoxic activity was tested against breast, melanoma-1, melanoma-2, and normal lymphoblast cells.

*1:4 dilution of concentration recommended by Wellcome Reagent Ltd. of PHA or PWM was used to stimulate lymphocytes. Con A was used at 1 μ g μ l⁻¹.

†A and A₁ are two different bleedings from donor 1.

on the melanoma-1 cells. Lymphocytes from two of three donors who had HL-A 2 were still immunised to melanoma-1 with bacterial extracts. This would suggest that the cytotoxicity observed was not directed towards the histocompatibility antigens of the target cells. The possibility that histocompatibility was not a factor in the phenomenon described will need to be confirmed in a syngeneic system in experimental animals.

In some experiments, cytotoxicity of bacterial immunised lymphocytes was greater than when lymphocytes were sensitised by tumour cells (Fig. 2). In addition, in a series of experiments, ⁵¹Cr melanoma-1 cells were incubated with rabbit antisera to BCG or LM for 30 min at 37 °C, washed twice, and then used as target cells. This further augmented the cytotoxicity of bacterial extract immunised lymphocytes (Fig. 2b). Normal sera from the same rabbits that had been obtained before immunisation did not augment cytotoxicity.

Our results indicate that some bacterial antigens can, in *in vitro* conditions, render lymphocytes from normal persons cytotoxic for some tumour cells. The observation that killing of tumour cells by immune lymphocytes was augmented if the tumour targets were coated with antibodies specific for the bacterial extract used suggests in addition that the antibody dependent cellular cytotoxicity (ADCC) mechanism may be involved. ADCC has been shown to be operative against a variety of tumour cells *in vitro*¹⁷.

The bacterial extracts used in this study are known to consist of many components, only a few of which may be critical for the effects described above. Further studies are needed to isolate and characterise the antigenic portions of bacterial extracts that transform lymphocytes from normal persons into lymphocytes with strong antitumour

activity. It is of interest in this regard that antigenic similarities between melanoma cells and BCG have been demonstrated^{8,9}. It is thus tempting to speculate that *in vivo* immunisation with bacteria which crossreacts with tumours may induce cytotoxic lymphocytes to certain tumour cells.

We thank Drs G. E. Moore and G. Seman for providing human tumour lines, and Anita Roybal, Kim Austin, Susan Beeler, Cynthia Bradley, and Pat Stem for technical assistance. This work was supported by the US NCI.

BRAHMA SHARMA*
DAVID G. TUBERGEN

The Children's Hospital,
Denver, Colorado 80218

PERCY MINDEN
MICHAEL J. BRUNDA

National Jewish Hospital and Research Center,
Denver, Colorado 80206

Received 8 February; accepted 26 April 1977.

*To whom reprint requests should be addressed. Present address: Department of Oncology-Hematology, The Children's Hospital, 1056 East Nineteenth Avenue, Denver, Colorado 80218.

- Sharma, B. & Terasaki, P. I. *Cancer Res.* 34, 115-118 (1974).
- Sharma, B. & Terasaki, P. I. *J. natn. Cancer Inst.* 52, 1925-1926 (1974).
- Sharma, B., Kohashi, O., Mickey, M. R. & Terasaki, P. I. *Cancer Res.* 35, 666-669 (1975).
- Sharma, B. *J. natn. Cancer Inst.* 57, 743-748 (1976).
- Golub, S. H. & Morton, D. L. *Nature* 251, 161-163 (1975).
- Borsos, T. & Rapp, H. J. *J. natn. Cancer Inst.* 51, 1085-1086 (1973).
- Minden, P., McClatchy, J. K., Wainberg, M. & Weiss, D. W. *J. natn. Cancer Inst.* 53, 1325-1331 (1974).
- Minden, P., Sharpton, T. R. & McClatchy, J. K. *J. Immun.* 116, 1407-1414 (1976).
- Bucana, C. & Hanna, M. G. *J. natn. Cancer Inst.* 53, 1313-1323 (1974).
- Minden, P., McClatchy, J. K. & Farr, R. S. *Infect. Immun.* 6, 574-582 (1972).
- Ferrari, A. *Ann N.Y. Acad. Sci.* 87, 792-800 (1960).
- Seman, G., Hunter, S. J., Miller, R. C. & Dmochowski, L. *Cancer* 37, 1814-1824 (1976).
- Holm, G., Perlman, P. & Werner, B. *Nature* 203, 841-843 (1964).
- Williams, T. W. & Granger, G. A. *Nature* 219, 1076-1077 (1968).
- Williams, T. W. & Granger, G. A. *Nature* 218, 1053-1055 (1968).
- Williams, T. W. & Granger, G. A. *J. Immun.* 102, 911-918 (1969).
- Sanderson, C. J., Clark, I. A. & Taylor, G. A. *Nature* 253, 376-377 (1975).

Cyclic nucleotide fluctuations during steroid induced meiotic maturation of frog oocytes

MATURING oocytes of the frog *Rana pipiens* arrest in prophase of the first meiotic division, when they can be isolated and induced to complete meiotic maturation *in vitro* by exposure to steroid hormone. A study of the meiotic maturation divisions of the frog oocyte therefore encompasses problems of steroid action and the control of cell division. An unusual feature of the interaction of steroids with the oocyte is the independence of the nucleus indicated by the failure of actinomycin D or enucleation to affect the induction of maturation by steroids¹. Furthermore, the interaction is nonspecific, in that several different steroids are capable of inducing maturation *in vitro*². Both cyclic AMP and theophylline inhibit progesterone induced maturation³. This report presents evidence that exposure of prophase-arrested oocytes to progesterone lowers intracellular cyclic AMP and cyclic GMP levels, that theophylline inhibits one component of this response, and that periodic cyclic nucleotide fluctuations occur during the subsequent maturation divisions.

The *R. pipiens* oocyte is particularly useful in an investigation of this nature. *In vitro* maturation of a group of isolated oocytes is intrinsically synchronous, and the appearance of characteristic morphological markers allows their progression through the meiotic divisions to be followed under the dissecting microscope. Furthermore, their extraordinarily large diameter—1,500 μm —allows microinjection of substances to active concentrations without markedly changing intracellular volume.

Gravid *R. pipiens* females obtained from December through April were maintained in artificial hibernation at 4 °C until sacrifice. Oocytes were manually dissected from their ovarian follicles with watchmaker's forceps, and exposed to progesterone at 5 mg l⁻¹ for 10 min. Incubations were carried out in amphibian Ringer's solution (100 mM NaCl, 1.9 mM KCl, 1.1 mM CaCl₂, 1.2 mM NaHCO₃, pH 7.8) at 22 °C. In control groups of 50–100 oocytes, at least 95% underwent proper development as determined by activation of the mature oocyte.

The cyclic nucleotide content of isolated *R. pipiens* oocytes which had not been exposed to progesterone (hereafter referred to as basal level) was 0.6 pmol cyclic AMP (range 0.3–1.1) and 8.0 pmol cyclic GMP (range 5.2–14.9) per oocyte, or approximately 0.6 μmol cyclic AMP and 8.0 μmol cyclic GMP per litre cell water (mean from ten experiments). The mean ratio of basal cyclic GMP content to cyclic AMP content was 15 (range 9–29), in marked contrast to the situation found in the oocytes of other amphibians⁴, and in other cell types (for example ref. 5). Basal levels found in samples of oocytes from the same female showed acceptable variation with standard errors which were 4–5% of the mean (three females, three samples per female, data not shown).

Time courses for fluctuations in the levels of cyclic AMP and cyclic GMP during the progesterone-induced *in vitro* maturation of isolated oocytes are presented in Fig. 1. The data show that significant decreases in the levels of both cyclic AMP and cyclic GMP occur within 30 min after exposure to progesterone. The decreases were not observed in experiments in which the control oocytes did not mature.

An antagonistic interaction between intracellular levels of cyclic AMP and cyclic GMP was formulated by Goldberg in terms of a yin-yang relationship¹¹, and McMahon has applied this concept to the regulation of cell division¹². Curiously, from the time of induction until germinal vesicle breakdown (GVBD), or dissolution of the prophase nucleus, the patterns of cyclic AMP and cyclic GMP

fluctuation are similar, yet after GVBD this trend is reversed and a reciprocal relationship is observed. According to Berridge¹³, these patterns of cyclic nucleotide fluctuations reflect monodirectional and bidirectional

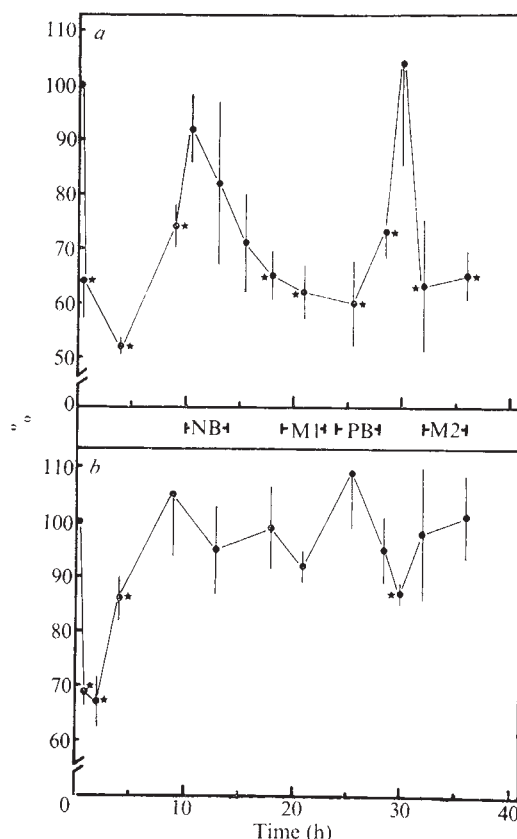


Fig. 1 Cyclic AMP (a) and cyclic GMP (b) levels during progesterone-induced maturation of isolated oocytes expressed as % of basal $\pm$ s.e.m. Abscissa shows time elapsed between exposure to progesterone and sampling. Zero on the time scale refers to basal cyclic nucleotide levels. At the times indicated, either a single sample of 30 oocytes, or duplicate samples of five oocytes were selected under the microscope on the basis of morphology when possible and analysed for cyclic nucleotide content. Cyclic nucleotides were extracted by rapidly homogenising samples of oocytes plus medium in two volumes of absolute ethanol at 0 °C. Immediately following homogenisation 3.8 nCi of ³H-cyclic AMP (38 Ci mmol⁻¹) and 2.1 nCi of ³H-cyclic GMP (3.5 Ci mmol⁻¹) were added to the homogenate in order to follow recovery from the purification steps. The tritiated cyclic nucleotides amounted to less than 5% and 4% of the total amount of cyclic AMP and cyclic GMP, respectively, in the sample. The homogenate was then centrifuged at 6,000 g for 10 min at 4 °C. The supernatant was flash evaporated under reduced pressure and frozen. Purification and separation of cyclic AMP and cyclic GMP was accomplished by sequential column chromatography over neutral alumina and Dowex 1 $\times$ 8 (formate form) according to Jakobs *et al.*⁶ and Butcher *et al.*⁷, respectively. Cyclic AMP and cyclic GMP were measured by the radioimmunoassay procedure of Steiner *et al.*⁸ as modified by Harper and Brooker⁹. Cyclic AMP was also measured by the protein binding assay of Brown *et al.*¹⁰ Both assay procedures for cyclic AMP gave the same results. Authenticity of cyclic nucleotides was established by incubating aliquots of purified extract with 10 μg of beef heart phosphodiesterase, after which less than 5% of the binding activity in the assays remained. Data from several experiments were pooled by first normalising the data within each experiment, expressing every point as % of basal level in oocytes from that particular female. Values from samples collected less than 90 min apart in different experiments were used to calculate a mean value of several experiments for cyclic nucleotide content at a given time point. All points are mean values $\pm$ s.e.m. from three or more experiments. Student's *t*-test was used to establish by paired comparisons whether there was a statistically significant difference between the mean at a given time point and the basal level (100%). Cyclic GMP was significantly different from basal at only one time point after GVBD. Stars indicate a statistically significant difference from basal ($P < 0.05$ – $P < 0.001$). NB, nuclear breakdown (GVBD); M1, first metaphase; PB, extrusion of first polar body; M2, arrest at second metaphase.

control, respectively. The apparent switchover from monodirectional to bidirectional control suggests that an alteration of cyclic nucleotide metabolism occurs at GVBD.

Berridge proposes that an antagonistic relationship exists between cyclic AMP and Ca^{2+} , with cyclic GMP levels simply reflecting intracellular Ca^{2+} levels. A number of studies have suggested that Ca^{2+} has an important role in the induction of maturation. It has been reported that alterations of calcium flux accompany progesterone induced maturation¹⁴, and that the divalent cation ionophore A23187 (ref. 15), as well as propanolol¹⁶, and lanthanum ions¹⁷ can induce maturation of *X. laevis* oocytes. Ionophore A23187 can induce maturation of *R. pipiens* oocytes (A. B. Kostellow, G. A. Morrill, personal communication), as well as parthenogenetically activate ova¹⁸.

Cleavage of the sea urchin egg is accompanied by a cycle of cyclic AMP fluctuations¹⁹ similar to those shown in Fig. 1, and similar fluctuations of cyclic AMP occur during proliferation of mammalian cells in culture²⁰. The significance of these fluctuations has not been established, however.

To investigate the physiological significance of decreased cyclic nucleotide levels in the mechanism of induction of maturation by progesterone, oocytes were exposed to the hormone in the presence of theophylline. Theophylline, a competitive inhibitor of cyclic nucleotide phosphodiesterase, was present at 2.0 mM during a 30 min preincubation, as well as during exposure to hormone, and continuously for the duration of the experiment. At this concentration, theophylline inhibited GVBD and subsequent maturation in 100% of the oocytes in control

groups. Theophylline inhibited the decrease in cyclic AMP levels, but in contrast did not inhibit the decrease in cyclic GMP levels after exposure to progesterone (Fig. 2).

The inhibitory effects of cyclic AMP and theophylline on maturation³ and of theophylline on the initial decrease in cyclic AMP levels after exposure to progesterone suggest that a decrease in the intracellular levels of this cyclic nucleotide is an early and necessary response of the oocyte in the induction of maturation by the steroid. Without characterising the oocyte adenylate cyclase and cyclic AMP phosphodiesterase(s), the data do not distinguish between activation of a phosphodiesterase and inhibition of adenylate cyclase as the event resulting in the observed decrease in cyclic AMP levels.

The failure of 2.0 mM theophylline to block the decline in cyclic GMP levels after hormone exposure suggests that this cyclic nucleotide is metabolised in a manner different from that of cyclic AMP. Alternatively, the subcellular distribution or kinetic properties of cyclic GMP phosphodiesterase different from those of cyclic AMP phosphodiesterase could be responsible for this result.

Cyclic nucleotide fluctuations are thought to influence cellular processes primarily through alterations in the activity of protein kinases^{21,22}. Increases in protein phosphorylation occur at GVBD in both *R. pipiens* and *X. laevis* oocytes^{23,24}. Maller and Krebs²⁵, using microinjection techniques, have shown that the regulatory subunit of protein kinase mimics progesterone by inducing maturation of *X. laevis* oocytes, while the catalytic subunit inhibits progesterone induced maturation. These results suggest that a decrease in the catalytic activity of oocyte protein kinase is a key step in the induction process in *X. laevis*, and are consistent with a decrease in cyclic AMP levels after exposure of *X. laevis* oocytes to progesterone, similar to that which we report here for *R. pipiens*.

This study of the isolated *R. pipiens* oocyte has revealed a novel mode of steroid action: progesterone causes a reduction in intracellular levels of cyclic AMP that is essential to release of the oocyte from prophase arrest. The significance of very high levels of cyclic GMP relative to cyclic AMP in these cells, and of the later cyclic nucleotide fluctuations have not been determined.

We thank Drs J. K. Reyhout and F. G. Rothman for assistance, Drs M. J. Birnbaum and F. T. Marin for discussions and criticism, and Drs A. B. Kostellow and G. A. Morrill for communication of unpublished results. This work was supported by the Cystic Fibrosis Foundation, the USPHS and NSF. F.R.B. is the recipient of a Research Career Development Award from the USPHS.

MARK G. SPEAKER*
FRED R. BUTCHER

Division of Biology and Medicine,
Brown University,
Providence, Rhode Island 02912

Received 27 December 1976; accepted 26 April 1977.

*To whom correspondence should be addressed at Box 209A, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461.

- Smith, L. D. & Ecker, R. E. *Dev. Biol.* **19**, 281-309 (1969).
- Schuetz, A. W. *Proc. Soc. Exp. Biol. Med.* **124**, 1307-1310 (1967).
- Schatz, F. & Morrill, G. A. *Fifth Int. Congr. Pharm.* (Abstract) (Karger, Basel, 1973).
- Pays-de Schutter, A., Kram, R., Hubert, E. & Brachet, J. *Exp. Cell Res.* **96**, 7-14 (1975).
- Pointer, R. H., Butcher, F. R. & Fain, J. N. *J. biol. Chem.* **251**, 2987-2992 (1976).
- Jakobs, K. H., Bohme, E. & Schultz, G. *Regulation of Function and Growth of Eukaryotic Cells by Intracellular Cyclic Nucleotides* (eds Dumont, J. E., Butcher, R. W. & Brown, G.) (Plenum, New York, 1976).
- Butcher, F. R., Rudich, L., Emiler, C. & Nemerovski, M. *Molec. Pharmacol.* **12**, 862-870 (1976).
- Steiner, A. L., Parker, C. W. & Kipnis, D. M. *J. biol. Chem.* **247**, 1106-1113 (1972).
- Harper, J. F. & Brooker, F. J. *cyclic Nucleotides Res.* **1**, 207-218 (1975).
- Brown, B. L., Albano, J. D. M., Elkins, R. P. & Szerzi, A. M. *Biochem. J.* **121**, 561-562 (1971).
- Goldberg, N. D., O'Dea, R. F. & Haddox, M. K. *Adv. cyclic Nucleotide Res.* **3**, 155-223 (1973).

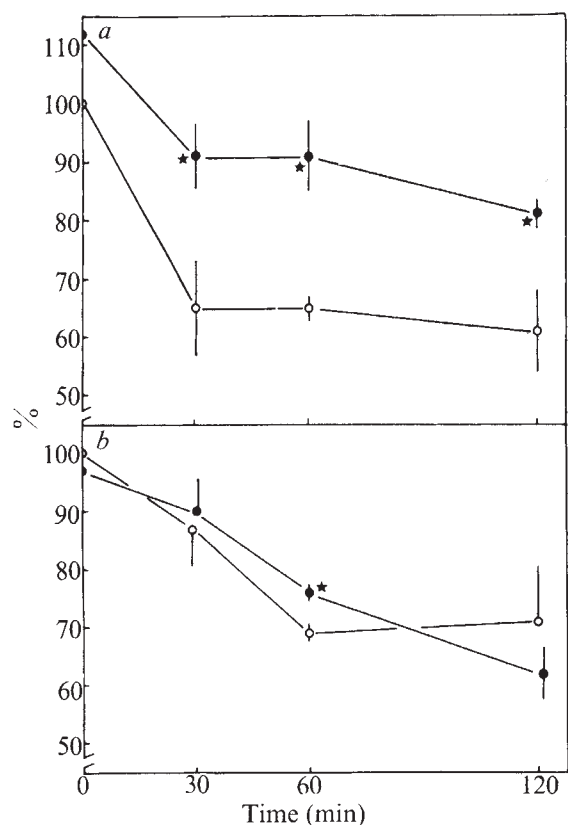


Fig. 2 Effect of theophylline on cyclic nucleotide fluctuations during the induction response. Cyclic nucleotide levels per oocyte expressed as % of basal $\pm$ s.e.m. All points are means from three experiments. Abscissa refers to time elapsed between exposure to progesterone and sampling. Stars indicate a statistically significant difference between control and + theophylline at a given time point ($P < 0.05$ to $P < 0.01$). a, Cyclic AMP levels and b, cyclic GMP levels. ○, Control; ●, plus 2.0 mM theophylline.

- ¹² McMahon, D. *Science* **185**, 1012-1021 (1974).
- ¹³ Berridge, M. J. *Adv. cyclic Nucleotide Res.* **6**, 1-98 (1975).
- ¹⁴ O'Connor, C. M. & Smith, L. D. *Dev. Biol.* **52**, 318-322 (1976).
- ¹⁵ Wasserman, W. J. & Masui, Y. *J. exp. Zool.* **193**, 369-375 (1975).
- ¹⁶ Schorderet-Slatkine, S. & Schorderet, M. *C. r. heb. Séanc. Acad. Sci., Paris* **282**, 1733-1736 (1976).
- ¹⁷ Schorderet-Slatkine, S., Schorderet, M. & Baulieu, E. E. *Nature* **262**, 289-290 (1976).
- ¹⁸ Schroeder, T. E. & Strickland, D. L. *Expl Cell Res.* **83**, 139-142 (1974).
- ¹⁹ Yasumasu, I., Fujiwara, A. & Ishida, K. *Biochem. biophys. Res. Commun.* **54**, 628-632 (1973).
- ²⁰ Pastan, I. H., Johnson, G. S. & Anderson, W. B. *A. Rev. Biochem.* **44**, 491-522 (1975).
- ²¹ Kuo, J. F. & Greengard, P. *J. biol. Chem.* **244**, 3417-3419 (1969).
- ²² Casnellie, J. E. & Greengard, P. *Proc. natn. Acad. Sci., U.S.A.* **71**, 1891-1895 (1974).
- ²³ Morrill, G. A. & Murphy, J. B. *Nature* **238**, 282-284 (1972).
- ²⁴ Maller, J., Wu, M. & Gerhart, J. C. *Dev. Biol.* (in the press).
- ²⁵ Maller, J. L. & Krebs, E. G. *J. biol. Chem.* **252**, 1712-1718 (1977).

Prostacyclin increases cyclic AMP levels and adenylate cyclase activity in platelets

PROSTACYCLIN (PGX) (ref. 1) is the name proposed for the metabolic product of a recently discovered enzymic transformation of the prostaglandin endoperoxides PGG_2 and PGH_2 (refs 2-4). Prostacyclin synthetase, the enzyme which catalyses this conversion, is particularly active in the microsomal fraction of blood vessel walls^{3,4} although there is evidence that it is also present in the rat stomach fundus⁵. The prostaglandin endoperoxides, PGG_2 and PGH_2 , have been shown to contract arterial smooth muscle and to cause platelet aggregation⁶, whereas, in contrast, prostacyclin relaxes arterial strips and prevents platelets aggregation³. A biochemical interaction has therefore been postulated to exist between platelets and the blood vessels wall, such that endoperoxides derived from platelets are converted by the vessel wall into prostacyclin, which in turn actively prevents the deposition of platelets on the vascular endothelium^{2,4}. In general, agents which inhibit platelet aggregation, such as PGE_1 , increase cyclic AMP levels, whereas agents causing aggregation lower platelet cyclic AMP levels⁷. The ability of prostacyclin to rapidly reverse or prevent platelet aggregation suggested that its action on the platelet might, like that of PGE_1 , be mediated by an increase in cyclic AMP levels. This hypothesis is supported by the experiments detailed here which strongly suggest that prostacyclin generated by incubation of arachidonic acid, platelets and aortic tissue, causes a rapid and pronounced accumulation of cyclic AMP by intact platelets, as well as stimulation of adenylate cyclase in isolated platelet membranes.

Human platelet-rich plasma (PRP) was incubated in the presence of theophylline, arachidonic acid, and chopped pieces of pig aorta obtained fresh from a local abattoir. Incubation conditions were similar to those used to generate prostacyclin in incubations of PRP with vascular rings⁴. Cyclic AMP was assayed in samples of incubation media at timed intervals. In the presence of arachidonic acid and aortic pieces, the cyclic AMP content of PRP rose approximately 12-fold within 5 min (Fig. 1a), declining gradually over the next 25 min. In the absence of aortic pieces, arachidonic acid produced no change in platelet cyclic AMP levels. No changes were observed in the cyclic AMP content of platelet-poor plasma (PPP) incubated in the presence of arachidonic acid and aortic rings. When arachidonic acid was omitted from the incubation medium, or the platelet-rich plasma was pre-incubated for 15 min with $100 \mu\text{g ml}^{-1}$ indomethacin, the changes in the cyclic AMP content of PRP were very much diminished, in one case presumably because substrate was absent, and in the other because the inhibition of cyclooxygenase by indomethacin led to reduced prostaglandin endoperoxide levels (Fig. 1b). In some experiments aortic pieces were incubated for 15 min with indomethacin and washed before incubation with PRP and arachidonic acid. This treatment did not alter the effect of aortic tissue on cyclic

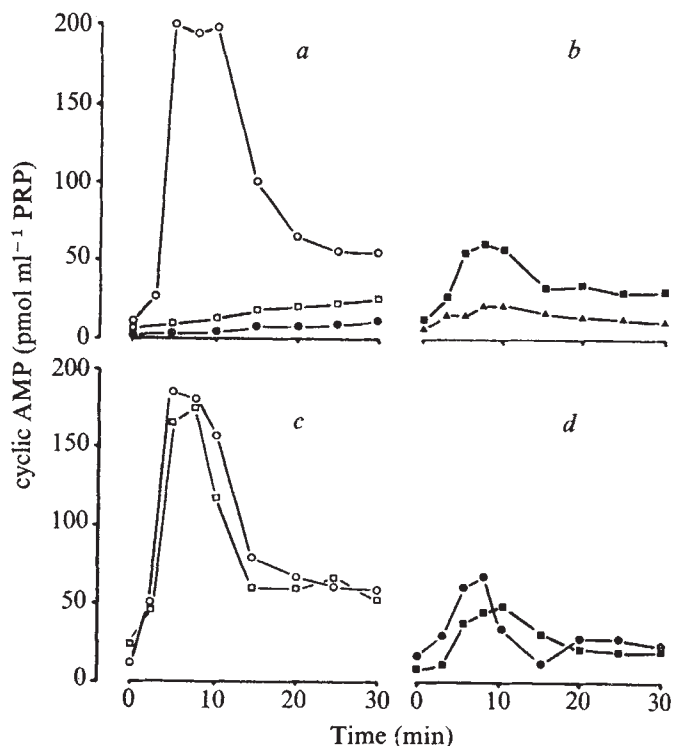


Fig. 1 Cyclic AMP generation by platelets from arachidonic acid in the presence and absence of aortic tissue. PRP was prepared by centrifuging venous blood (9 volumes: 1 volume 3.8% citrate) at 500 *g* for 10 min at room temperature; PPP was obtained by centrifuging PRP at 5,000 *g* for 10 min. Reactions were begun by adding 4 ml of PRP or PPP to incubation flasks containing 2 ml buffer (15 mM Tris, pH 7.4, with 120 mM NaCl, 4 mM KCl, 1.6 mM MgSO_4 , 2 mM NaH_2PO_4 , 10 mM glucose and 0.2% bovine serum albumin) and theophylline (10 mM final concentration). Final concentrations of other components of the incubation medium were as follows: fresh pig aortic pieces (50 mg ml^{-1}), arachidonic acid (Sigma; AA, 10 $\mu\text{g ml}^{-1}$). Where stated, PRP was pre-incubated for 15 min with indomethacin (I; 10 $\mu\text{g ml}^{-1}$). 500 μl samples of incubation media were taken at intervals, boiled for 2 min, and the cyclic AMP content measured using beef adrenal protein kinase competitive binding assay⁸. Standard curves included an appropriate amount of a mixture of buffer and PPP depleted of cyclic AMP by standing overnight at room temperature. Assays were carried out in triplicate on single samples obtained at the indicated time points. The cyclic AMP content of the mixture of buffer and PPP was assayed and the results corrected for this value. *a*, PRP+AA and aortic pieces (○); PRP+AA (□); PPP+AA+aortic pieces (●). *b*, PRP+aortic pieces (■); PRP (+I)+AA+aortic pieces (▲). *c*, PRP+AA+aortic pieces+U44069 (○); PRP+AA+aortic pieces+U46619 (□). *d*, PRP+aortic pieces+U44069 (●); PRP+aortic pieces+U46619 (■).

AMP levels (results not shown) and is consistent with the reported resistance of prostacyclin synthetase to inhibition by indomethacin. These results are consistent with the view that prostaglandin endoperoxides, intermediates in the metabolism of arachidonic acid by platelets, were converted in the presence of the aortic pieces to a substance which caused an immediate, pronounced rise in platelet cyclic AMP content. The same conditions, used by Bunting *et al.*⁴, resulted in formation of prostacyclin, a potent inhibitor of platelet aggregation.

The prostaglandin endoperoxide analogues (15*S*)-hydroxy-9 α , 11 α (epoxymethano) prosta 5*Z*-13*E*-dienoic acid (U44069) and (15*S*)-hydroxy-11 α , 9 α (epoxymethano) prosta 5*Z*-13*E* diennoic acid (U46619) had no influence on prostacyclin formation during the incubation of PRP with aortic rings (Fig. 1c, d), suggesting that as expected, these analogues were not used as substrates by prostacyclin synthetase and neither had any inhibitory effect on the enzyme. But, the addition of an aggregating agent—either

ADP or adrenaline—caused a marked decline in prostacyclin-stimulated platelet cyclic AMP levels (Fig. 2a). This effect was comparable to that of ADP or adrenaline on platelet cyclic AMP levels previously elevated by treatment with prostaglandin E_1 , and the mechanism may be that ADP and adrenaline directly stimulate phosphodiesterase in platelet membranes⁹. The stimulation of cyclic AMP formation during incubation of PRP with arachidonic acid and aortic fragments was prevented by including tranilcypromine (500 μ M) or oxidised arachidonic acid (10 μ g ml^{-1}) (Fig. 2b). Each of these substances has been shown to inhibit the activity of prostacyclin synthetase³.

A sample of incubation mixture of PRP, arachidonic acid and aortic pieces stimulated the activity of adenylate cyclase in partially purified platelet membranes (Fig. 3). Thus, the elevation by prostacyclin of platelet cyclic AMP levels, via the activation of platelet membrane adenylate cyclase, resembles the effects of PGE_1 and may account for the anti-aggregating property of prostacyclin previously demonstrated by Gryglewski *et al.*³. This conclusion receives support now from the demonstration that chemically synthesised prostacyclin is many times more active than PGD_2 and PGE_1 in stimulating platelet production of cyclic AMP and platelet membrane adenylate cyclase^{12,13}.

Involvement of cyclic AMP as an intermediate in prostacyclin action provides a biochemical basis for an important

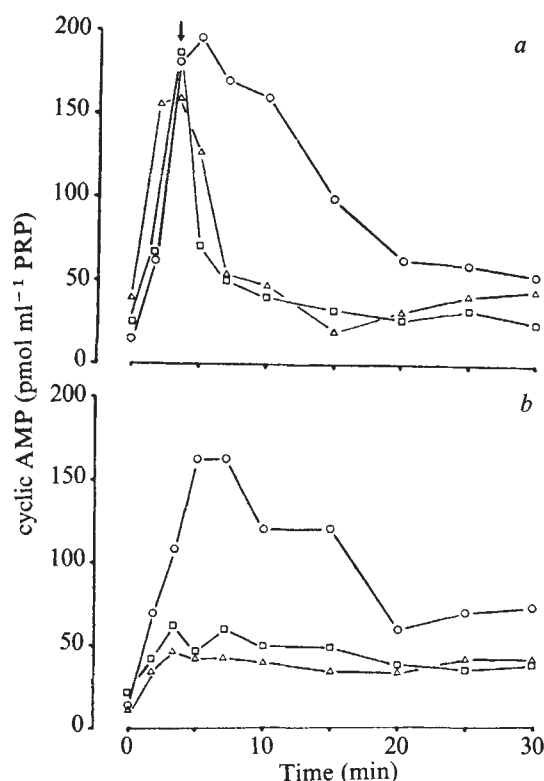


Fig. 2a Effects of ADP and adrenaline on accumulation of cyclic AMP by platelets. Incubations were carried out as described for Fig. 1, and contained PRP in buffer, AA (10 μ g ml^{-1}) and aortic rings (50 mg ml^{-1}). At the time indicated by the arrow, ADP (10 μ g ml^{-1} ; □) or adrenaline (10 μ g ml^{-1} ; △) or buffer (O) was added, and sampling for cyclic AMP assay continued. **b**, Effects of tranilcypromine and of oxidised arachidonic acid on cyclic AMP accumulation by platelets in the presence of aortic pieces. Arachidonic acid was oxidised by dissolving it in buffer at 10 μ g ml^{-1} and exposing it to a stream of air overnight at 25 °C. Incubations were carried out as described in Fig. 1. PRP + aortic pieces + AA (O); PRP + aortic pieces + AA + tranilcypromine (0.5 mM) (□); PRP + aortic pieces + AA + oxidised arachidonic acid (10 μ g ml^{-1}) (△). Assays were carried out in triplicate on single samples obtained at the indicated time points.

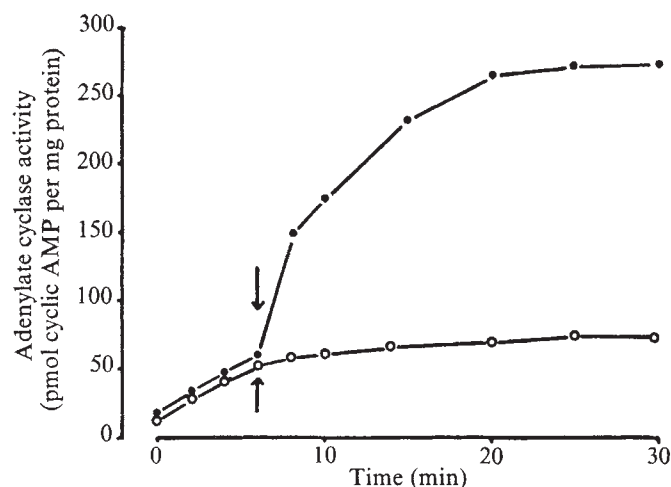


Fig. 3 Adenylate cyclase activity in crude platelet plasma membranes. Platelets were centrifuged from PRP at 5,000 g for 10 min, the platelet pellet disrupted by rapid freezing and thawing (three times) and the crude membrane fraction harvested by centrifuging at 10,000 g . Adenylate cyclase incubations were carried out in 2 ml volumes of 25 mM Tris HCl containing 1 mM ATP³² (10 c.p.m. $pmol^{-1}$), 4.5 mM $MgCl_2$, 30 mM KCl, 1.5 IU pyruvate kinase; 2 mM phosphoenolpyruvate; 0.013% bovine serum albumin, and platelet membranes. After starting the reactions by addition of platelet membranes, 100- μ l samples were removed at the intervals shown, added to 10 μ l of 40 mM ATP and boiled immediately. The generated ³²P-cyclic AMP was purified by sequential chromatography on Dowex and alumina^{10,11}. At the time indicated by the arrow further additions were made to each flask as follows: ●, 100 μ l from a 6-min incubation of PRP, AA (10 μ g ml^{-1}) and aortic rings (50 mg ml^{-1}) which had been incubating for 5 min (see Fig. 1a). ○, 100 μ l from a 6-min incubation of PRP and AA (10 μ g ml^{-1}) which had been incubating for 5 min (see Fig. 1a).

cellular control system, which produces cooperation between the vascular wall and the blood platelet in order to prevent platelet adhesion to vascular endothelium. In this homeostatic mechanism, the prostacyclin could act continuously upon the platelets to increase cyclic AMP, which in turn may inhibit the platelet cyclooxygenase¹⁴, thereby limiting the amount of endoperoxide available for thromboxane synthesis, and thus turning off production of these potent short-lived stimulators of platelet aggregation. If prostacyclin or related compounds are found to be produced in other tissues, studies of their effects on target tissues will be greatly facilitated if they are found to have the general property of stimulating adenylate cyclase activity.

The support of Boehringer Ingelheim is acknowledged, and we thank Dr J. E. Pike (Upjohn) for prostaglandins and analogues.

L. C. BEST
T. J. MARTIN
R. G. G. RUSSELL
F. E. PRESTON

Departments of Chemical Pathology
and Haematology,
Academic Division of Pathology,
University of Sheffield Medical School,
Sheffield, UK

Received 21 March; accepted 6 May 1977.

- Johnson, R. A. *et al. Prostaglandins* 12, 915–928 (1976).
- Moncada, S., Gryglewski, R., Bunting, S. & Vane, J. R. *Nature*, 263, 663–665 (1976).
- Gryglewski, R. J., Bunting, S., Moncada, S., Flower, R. J. & Vane, J. R. *Prostaglandins* 12, 685–713 (1976).
- Bunting, S., Gryglewski, R., Moncada, S. & Vane, J. R. *Prostaglandins* 12, 897–913 (1976).

- ⁵ Hamberg, M. & Samuelsson, B. *Proc. natn. Acad. Sci. U.S.A.* **70**, 899–903 (1973).
⁶ Svensson, J., Hamberg, M. & Samuelsson, B. *Acta physiol. scand.* **94**, 222–228 (1975).
⁷ Mills, D. C. B. & Smith, J. B. *Biochem. J.* **121**, 185–196 (1971).
⁸ Brown, B. L., Albana, J. D. M., Ekins, R. P. & Sgherzi, A. M. *Biochem. J.* **121**, 561–562 (1971).
⁹ Best, L. C., Russell, R. G. G., Preston, F. E. & Martin, T. J. J. *Endocr.* (in the press).
¹⁰ Salomon, Y., Londos, C. & Rodbell, M. *Analyt. Biochem.* **58**, 541–548 (1974).
¹¹ Hunt, N. H., Martin, T. J., Michelangeli, V. P. & Eisman, J. A. *J. Endocr.* **69**, 401–412 (1976).
¹² Tateoson, J. E., Moncada, S. & Vane, J. R. *Prostaglandins* **13**, 389–397 (1977).
¹³ Gorman, R. R., Bunting, S. & Miller, O. V. *Prostaglandins* **13**, 377–388 (1977).
¹⁴ Malmsten, C., Granström, E. & Samuelsson, B. *Biochem. biophys. Res. Commun.* **68**, 569–576 (1976).

Calcium-dependent somatostatin secretion from rat neurohypophysis *in vitro*

THE tetradecapeptide somatostatin originally identified in hypothalamic extracts, has been subsequently shown by bioassay, radioimmunoassay and immunohistochemical techniques to be extensively distributed throughout the central nervous system, in the gastrointestinal tract, and in the pancreas^{1–3}. Within the central nervous system, immunoreactive somatostatin has been localised to fibres in the external zone of the median eminence, to neuronal cell bodies in the anterior hypothalamus, to axons in the posterior pituitary lobe, as well as in spinal ganglia and in fibres in the dorsal zone of the spinal cord^{2,6}. Although strong indirect evidence exists for the release of somatostatin from the hypothalamus into the portal vessels in physiological conditions⁷ there are no direct studies of somatostatin secretion from nerve cells. Since the neurohypophysis contains a high concentration of somatostatin⁸, and can be readily isolated, we used it as a model for studying the *in vitro* release of the peptide from nerve terminals. We report that secretion is markedly increased following potassium stimulation, provided calcium is present in the incubation medium.

Male Sprague–Dawley rats (240–300 g) were killed by decapitation, the posterior pituitary lobes were isolated, separated from adhering intermediate lobe tissues and incubated in pools of five in 1 ml bicarbonate-buffered Locke's solution at 37 °C under an atmosphere of 95% O₂–5% CO₂ with constant shaking. After a 30 min preincubation period, the medium was replaced by fresh Locke's solution or a modified Locke's solution at 30 min intervals. Medium from each incubation period was collected in glass tubes at 0 °C. On completion of each experiment the neural lobes were extracted by sonication in 1 ml 1-M acetic acid. Media and extracts were measured for somatostatin by a specific radioimmunoassay⁸ with a mean sensitivity of 1.4 pg. The protein content of the extracts was determined by the method of Lowry *et al.*⁹.

A high concentration of immunoreactive somatostatin was found in the posterior pituitary extracts. This contrasts with the absence of somatostatin in the anterior pituitary⁸. The total neurohypophysial content of somatostatin was determined by adding the amount of immunoreactivity released in the incubation media to the content remaining in the neurohypophysial lobes at the end of the experiment and yielded a value of 773±64 pg per neural lobe (mean±s.e.m. of 84 pools of five lobes each) or a concentration of 4.9±0.4 pg per µg protein. Somatostatin was secreted at a low basal rate from the isolated neural lobes, the amount released 5.9±1.4 pg per neural lobe per 30 min, representing 1.4±0.35% of the total content (mean of 20 experiments, Fig. 1).

High extracellular potassium is a potent stimulus for neurohypophysial hormone secretion and this response can be potentiated by a lowering of the extracellular sodium ion concentration¹⁰. The release of somatostatin was also markedly increased following potassium stimulation to 14.8% of total content per 30 min ($P<0.01$, Fig. 1a). This stimulatory effect was calcium dependent and was abolished

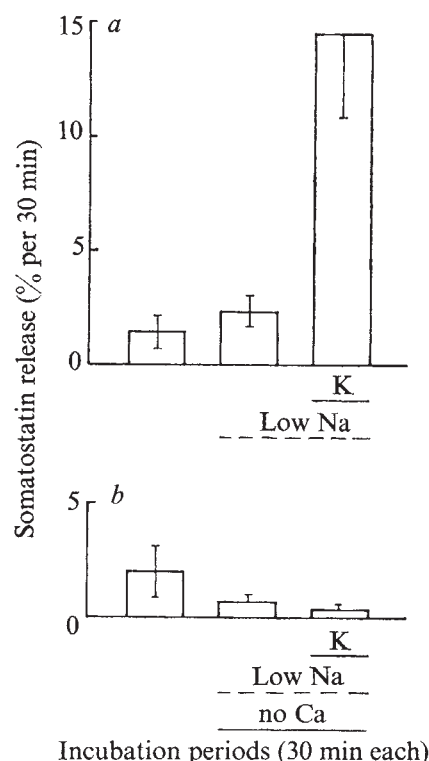


Fig. 1 Somatostatin-like immunoreactivity released by pools of five rat neural lobes during successive 30 min incubation periods. Each column represents the mean ± s.e.m. of 10 experiments. The initial incubation medium (first column in a and b) contained (in mM): Na⁺, 160; K⁺, 5.6; Ca²⁺, 2.2; Mg²⁺, 1.0; Cl⁻, 160; HCO₃⁻, 12 and glucose, 10. In a, the Na⁺-concentration of incubation medium was lowered to 12 mM for the second and third incubation periods; in addition, the K⁺-concentration was raised to 56 mM during the third period. Isoosmolarity was maintained by adding appropriate amounts of choline. In b, the same protocol was followed except that Ca²⁺ was replaced by Mn²⁺ (2.2 mM) during the second and third incubations. Values are expressed in % of total content of neurohypophysial somatostatin released per incubation period.

when extracellular Ca²⁺ was replaced by an equimolar concentration of Mn²⁺ an antagonist of Ca²⁺ in secretory processes (Fig. 1b, $P<0.01$).

The present studies confirm our earlier observation of the presence of immunoreactive somatostatin in the rat neurohypophysis⁸. Somatostatin thus represents a further addition to the list of peptides found in the posterior pituitary which includes oxytocin, vasopressin and thyrotrophin releasing hormone¹¹. Although the posterior pituitary concentration of somatostatin is comparable with that of other somatostatin rich neural tissues (for example, cerebral cortex, brain stem, spinal cord⁸) it is over 1,000 times lower than the concentration of the classical neurohypophysial hormones, oxytocin and vasopressin^{12,13}.

The mechanism of somatostatin release from the posterior pituitary shares two important characteristics with the neurosecretory process: stimulation of release in response to membrane depolarisation and dependence of this process on extracellular calcium. The functional significance of neurohypophysial somatostatin remains to be determined. One possibility is that somatostatin stimulates the release of vasopressin as suggested by *in vitro* studies¹⁴. Second, since somatostatin released in the posterior pituitary is capable of reaching the anterior pituitary through the short portal vessels, it seems likely that neurohypophysial somatostatin along with somatostatin released from the median eminence may have a role in the regulation of growth hormone and thyroid stimulating hormone secretion.

We thank Miss S. J. Franklin for technical assistance. This work was supported by the National Health and MRC

of Australia and by the Swiss NSF.

Y. C. PATEL

Medical Research Centre,
Prince Henry's Hospital,
Melbourne, Australia, and
Institute of Histology,
The Medical School,
Geneva, Switzerland

H. H. ZINGG
J. J. DREIFUSS

Department of Physiology,
The Medical School,
Geneva, Switzerland

Received 24 January; accepted 23 February 1977.

- ¹ Vale, W. *et al. Rec. Progr. Horm. Res.* 31, 365 (1975).
- ² Hökfelt, T. *et al. Acta endocrinol.* 80, 5-41 (1975).
- ³ Dubois, M. P. *Proc. natn. Acad. Sci. U.S.A.* 72, 1340-1343 (1975).
- ⁴ Brownstein, M., Arimura, A., Sato, H., Schally, A. V. & Kizer, J. S. *Endocrinology* 96, 1456-1461 (1975).
- ⁵ Patel, Y. C., Rao, K. & Reichlin, S. *New Eng. J. Med.* (in the press).
- ⁶ Alpert, L., Braver, J. R., Patel, Y. C. & Reichlin, S. *Endocrinology* 98, 255-258 (1976).
- ⁷ Terry, L. O., Willoughby, J. O., Brazeau, P., Martin, J. B. & Patel, Y. C. *Science* 192, 565-567 (1976).
- ⁸ Patel, Y. C. & Reichlin, S. in *Methods of Hormone Radioimmunoassay* (eds Jaffe, B. M. & Behrman, H. R.) (Academic, New York, in the press).
- ⁹ Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *biol. Chem.* 193, 265-271 (1951).
- ¹⁰ Dreifuss, J. J., Grau, J. D. & Bianchi, R. *Experientia* 27, 1295-1296 (1971).
- ¹¹ Oliver, C., Eskay, R. L., Ben-Jonathan N. & Porter J. C. *Endocrinology* 95, 540-546 (1974).
- ¹² George J. M., Staples S. & Marks, B. M. *Endocrinology* 98, 1430-1433 (1976).
- ¹³ George, J. M. *Science* 193, 146-148 (1976).
- ¹⁴ Skowsky, R. & Swan, L. *Proc. 5th International Congr. Endocrinol. Hamburg, July, 1976, 197 (abstract).*

Brain thyrotrophin-releasing hormone is independent of the hypothalamus

THYROTROPHIN-releasing hormone (TRH) is widely distributed in regions of the brain outside the hypothalamus¹⁻³, and behavioural⁴ and neurophysiological⁵ studies lend credence to the hypothesis^{1,6} that in these locations it is involved in neurotransmission. To determine whether extrahypothalamic brain TRH arises from the hypothalamus or is synthesised *in situ*, we have made electrolytic lesions of the "thyrotrophic area"⁷ of the rat hypothalamus, and examined the effect on TRH content in the hypothalamus itself, the rest of the brain and the anterior and posterior pituitary. We have found that such lesions markedly reduce the hypothalamic content of TRH, deplete the pituitary gland of TRH and result in severe chemical hypothyroidism, but do not affect the large amount of TRH normally stored in the extrahypothalamic brain.

We placed bilateral electrolytic hypothalamic lesions with a platinum electrode (Stoelting lesion maker) (3mA for 30 s) in male Sprague-Dawley rats (Charles River) under ether anaesthesia. The coordinates⁸ chosen—AP 6.4 mm, lateral 0.7 mm on each side of midline, vertical +1.5 mm—were designed to produce a large lesion in the region of the paraventricular nucleus extending towards, but not involving, the median eminence or pituitary stalk. Such lesions produce marked impairment of pituitary-thyroid function⁹. Six out of nine animals survived the procedure. In the seven control rats bilateral electrolytic lesions were placed in the cerebral cortex 2 mm below the surface. Both groups weighed about 400 g.

The effectiveness of the procedure was assessed 1 week later by determination of serum thyroxine (T4) by immunoassay on blood removed by orbital puncture from the rats under ether anaesthesia. T4 levels in the lesioned rats were all lower than the values in the controls (mean $\pm$ s.e.m.) 0.7 ± 0.1 compared with 4.1 ± 0.4 $\mu\text{g dl}^{-1}$ ($P < 0.001$).

To obviate the effect of hypothyroidism *per se* on tissue TRH content, all animals (test and control) were given orally sodium-L-thyroxine ($12 \mu\text{g d}^{-1}$, a dose designed to provide normal thyroid replacement) in 20 ml of a synthetic liquid diet (Sustacal) for 14 d, the last dose of thyroxine being given 48 h before termination of the experiment. The animals were decapitated and the brains were removed immediately. The hypothalamic lesions were confirmed visually with the aid of a dissecting

microscope, and hypothalami from both groups were excised and placed in 2 ml of 90% methanol for extraction; the anterior and posterior pituitaries were treated similarly. Tissue taken for the hypothalamic fragment was bounded anteriorly by the optic chiasm, posteriorly by the anterior margin of the mammillary nucleus, laterally by the hypothalamic sulcus and dorsally by a plane connecting the anterior commissure and anterior margin of the mammillary nucleus. To extract TRH from the brain, it was extruded through a 21-gauge needle into 10 ml of 90% methanol. The dried methanol extracts of all the tissues were reconstituted in phospho-saline buffer and assayed for TRH by a specific radioimmunoassay as before¹.

Although the hypothalamic content of TRH was much reduced in the lesioned rats compared with controls ($3,625 \pm 249$ compared with $9,357 \pm 449$ pg; $P < 0.001$), the amount of TRH in the extrahypothalamic brain ($17,040 \pm 925$ compared with $18,929 \pm 737$ pg) was similar in both groups; indeed the concentration of TRH was identical in the lesioned and normal animals (70 pg per mg protein¹⁰). In contrast the TRH content was markedly lower in the anterior pituitary (15 ± 8 compared with 64 ± 5 pg; $P < 0.001$) and posterior pituitary (< 2 compared with 157 ± 32 pg; $P < 0.001$) (Table 1). The lower TRH in anterior pituitary likely resulted because less TRH was delivered to the pituitary.

Table 1 Effect of a lesion of the 'thyrotrophic area' of the hypothalamus on brain distribution of TRH in the rat

	Lesion	Control	Significance
Hypothalamus	$3,625 \pm 249$	$9,357 \pm 449$	$P < 0.001$
Anterior pituitary	15 ± 8	64 ± 5	$P < 0.001$
Posterior pituitary	< 2	157 ± 32	$P < 0.001$
Extrahypothalamic brain	$17,040 \pm 925$	$18,929 \pm 737$	$P > 0.1$

Results (mean $\pm$ s.e.m.) are given as pg per tissue. Six lesioned and seven control animals were studied.

This study shows that a lesion of the hypothalamus resulting in reduced hypothalamic content of TRH and impaired TRH secretion from the hypothalamus, as implied from the low circulating levels of thyroxine, does not affect extrahypothalamic brain TRH content, suggesting that the latter is synthesised *in situ* and is independent of the hypothalamic secretion. Complementary studies were done by Brownstein *et al.*¹¹ who showed that after hypothalamic deafferentation the extrahypothalamic brain TRH content was unaltered whereas the hypothalamic TRH content was reduced, raising the possibility that much of the hypothalamic TRH is synthesised by cells outside this region.

The high concentration of TRH in the posterior pituitary relative to the anterior pituitary in the normal group (560 ± 9 compared with 28 ± 2 pg per mg protein; $P < 0.001$) has been observed before^{1,3}. The depletion of TRH from the posterior pituitary in the lesioned animals supports the concept of a third neurosecretory hypothalamo-hypophysial system extending into the neurohypophysis, as suggested by Hökfelt *et al.*¹² on the basis of immunohistochemical staining of TRH¹³ and somatostatin¹³ positive fibres reaching into the posterior pituitary from the median eminence. Whether TRH has a role in posterior pituitary function is uncertain, but it may be relevant that TRH is found in high concentrations in the pituitary complex of lower vertebrates¹, and there is evidence (reviewed by Sawyer¹⁴) that in the bony fish the neurohypophysis may be an homologue of the median eminence in higher animals.

We thank Mrs Faye SooHoo for technical assistance. This study was supported by a grant from the NIH.

IVOR M. D. JACKSON
SEYMOUR REICHLIN

Endocrine Division,
Department of Medicine,
New England Medical Center Hospital and
Tufts University School of Medicine,
Boston, Massachusetts 02111

Received 21 March; accepted 28 April 1977.

- ¹ Jackson, I. M. D. & Reichlin, S. *Endocrinology* **95**, 854–862 (1974).
- ² Winokur, A. & Utiger, R. D. *Science* **185**, 265–267 (1974).
- ³ Oliver, C., Eskay, R. L., Ben-Jonathan, N. & Porter, J. C. *Endocrinology* **95**, 540–546 (1974).
- ⁴ Plotnikoff, N. P., Prange Jr, A. J., Breese, G. R., Anderson, M. S. & Wilson, I. C. *Science* **178**, 417–418 (1972).
- ⁵ Renaud, L. P. & Martin, J. B. *Brain Res.* **86**, 150–154 (1975).
- ⁶ Brownstein, M. J., Palkovits, M., Saavedra, J. M., Bassiri, R. M. & Utiger, R. D. *Endocrinology* **185**, 267–269 (1974).
- ⁷ Halász, B., Pupp, L. & Uhlarik, S. *J. Endocr.* **25**, 147–154 (1962).
- ⁸ Albe-Fessard, D., Stutinsky, F. & Libouban, S. *Atlas Stéréotaxique du Diencephale du Rat Blanc* (Editions du Centre National de la Recherche Scientifique, Paris, 1966).
- ⁹ Martin, J. B., Boshans, R. & Reichlin, S. *Endocrinology* **87**, 1032–1040 (1970).
- ¹⁰ Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
- ¹¹ Brownstein, M. J., Utiger, R. D., Palkovits, M. & Kizer, J. S. *Proc. Natn. Acad. Sci. U.S.A.* **72**, 4177–4179 (1975).
- ¹² Hökfelt, T. *et al.*, *Acta Endocr. Suppl.*, 5–41 (1975).
- ¹³ Hökfelt, T., Fuxe, K., Johansson, L., Jeffcoate, S. & White, N. *Eur. J. Pharmac.* **34**, 389–392 (1975).
- ¹⁴ Sawyer, W. H. *Rec. Prog. Hormone Res.* **17**, 437–465 (1961).

Synaptic responses mediated by identified histamine-containing neurones

EFFORTS to prove that histamine has a role in synaptic neurotransmission have been less successful than for other biogenic amines. For example, there has been no convincing physiological or pharmacological evidence to support the transmitter candidacy of this imidazole at any known synapse. In the central nervous system (CNS) of the marine mollusc *Aplysia californica* chemical measurements of single isolated neurones have revealed large concentrations of histamine within two identified neurones¹. The histamine-containing neurones (HN) uniquely possess a specific histidine decarboxylating enzyme and they can synthesise and store labelled histamine². These biochemical observations prompted the suggestion³ that the two identified nerve cells utilise histamine as a neurotransmitter. The physiological studies reported here show that each HN monosynaptically elicits multicomponent excitatory and inhibitory postsynaptic potentials (p.s.ps) in different follower neurones, the somal membranes of which show appropriate potential changes to iontophoretic application of histamine. This suggests that histamine is a neurotransmitter released from the nerve terminals of the HN.

In the absence of ultrastructural evidence needed to prove the monosynaptic nature of cell-to-cell connectivity, several physiological criteria provide a reliable basis on

which to distinguish between direct and indirect synaptic connections (reviewed in ref. 3). These criteria include: (1) that the postsynaptic response follows the presynaptic action potential in a one-for-one manner and that there is a constant latency of response; (2) that the postsynaptic responses persist in medium containing increased concentrations of Ca^{2+} , and (3) that synaptic efficacy increases after intracellular injection of tetraethylammonium (TEA) into the presynaptic neurone. These tests have been applied in the work reported here to support the monosynaptic character of synaptic responses elicited by the HN.

Desheathed cerebral ganglia dissected from *Aplysia* (200–350 g) were perfused with artificial seawater (ASW, 492 mM, NaCl; 10 mM, KCl; 20 mM, MgCl_2 ; 30 mM, MgSO_4 ; 60 mM, CaCl_2 ; 10 mM, Tris). The pH of the ASW was adjusted to 7.8 and the temperature was maintained at 18 °C by a Peltier device. Standard electrophysiological techniques were used for intracellular stimulation and recording with double-barrelled glass microelectrodes filled with 2 M potassium citrate or 3 M TEA bromide. In several experiments the identity of the HN was biochemically confirmed after intracellular recording from pre- and post-synaptic neurones by assaying the activity of histidine decarboxylase, which is a cell-specific marker for HN in *Aplysia*².

In either normal or Ca^{2+} -enriched ASW solutions the HN are silent, displaying little spontaneous synaptic and action potential activity. Several identifiable neurones located on the dorsal surface of the ganglion showed consistent synaptic effects after action potentials evoked in the HN. These follower neurones are symmetrically located on both the left and right sides of the ganglion: they show (1) identical responses to ipsilateral HN stimulation, (2) similar spontaneous synaptic and action potential rhythms and (3) antidromic responses to electrical stimulation of the corresponding ipsilateral and contralateral cerebral nerve trunks. No synaptic responses could be observed between left and right HN nor between ipsilateral HN and contralateral follower neurones; intracellular dye (ccbalt) injection also indicates that neither the HN nor the follower neurones send axons across the midline of the ganglion (my unpublished observations).

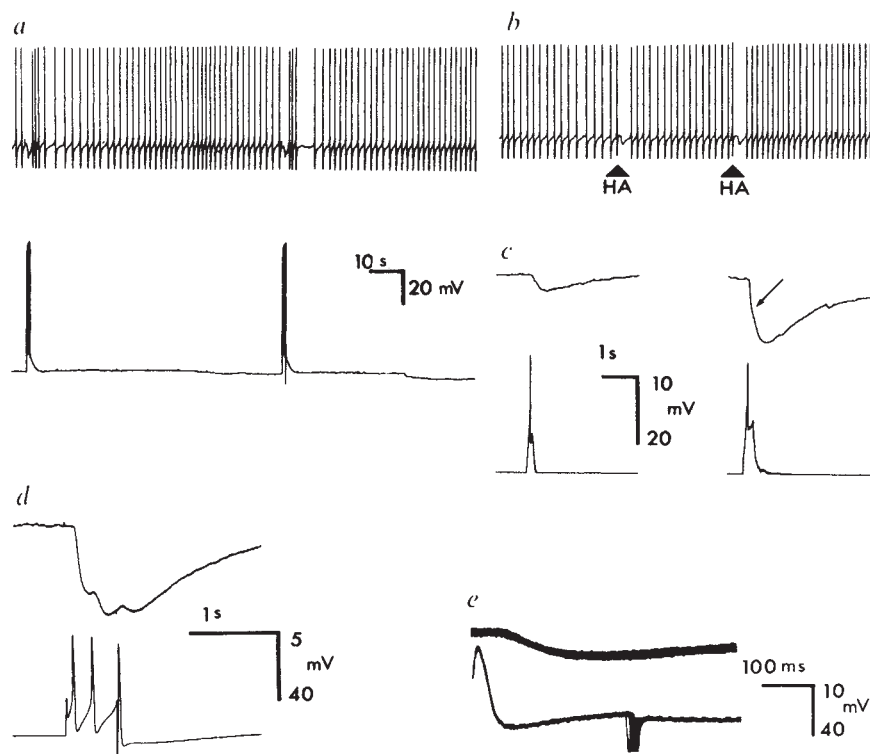


Fig. 1 Inhibitory effects produced by histamine-containing neurone (HN) spikes on one class of follower neurones. *a*, Inhibition of spontaneous spike activity in a follower neurone evoked by HN spikes. *b*, Inhibitory response to iontophoretic application (30 nA, 0.5 s) of histamine (HA) delivered to the somal membrane of the same follower cell; action potential amplitudes smaller than in (*a*) due to frequency response characteristics of the recording channel. *c*, i.p.s.ps recorded after spontaneous spiking was prevented by pre-setting the membrane potential to -55 mV. Right trace shows i.p.s.p. 5 min after TEA was injected into the HN by passing current between two intracellular TEA electrodes. Inflection point on the rising phase (arrow) of the i.p.s.p. depicts the fast phase of the two-component i.p.s.p. *d* and *e*, Fast i.p.s.ps recorded in a similar follower neurone from another cerebral ganglion at 4 °C to abolish the slow i.p.s.p.; membrane potential preset to about -50 mV. *d*, Association between HN spikes and fast i.p.s.ps; *e*, six superimposed records of fast i.p.s.ps and HN spikes illustrating constant latency of the evoked i.p.s.p.; latency of fast i.p.s.p. exaggerated by cooling. In this and subsequent figures upper voltage calibrations are associated with upper traces (follower neurones); lower voltage calibration corresponds with lower traces (HN).

Figure 1a shows the effects produced by a train of HN spikes on the endogenous activity of one class of follower neurones. At the normal resting potential, about -50 mV, as few as five HN spikes rapidly depress spontaneous spiking activity in these neurones. Thus the influence of the HN on these follower cells is clearly inhibitory. Spiking activity in the follower neurones was also inhibited by iontophoretic application of histamine (Fig. 1b), suggesting that histamine receptors are present on these postsynaptic membranes. When the membrane potential of the follower neurone was hyperpolarised beyond spiking threshold, hyperpolarising p.s.ps (i.p.s.ps) were observed. These are composed of two components, an initial rapid i.p.s.p. with a fast rise time and short duration (200–400 ms) and a slow i.p.s.p. with a gradual rise time and a prolonged decay phase (several seconds). The amplitude of the fast i.p.s.p. is usually small at membrane potentials near -60 mV and frequently obscured by the slow i.p.s.p. The fast i.p.s.p. becomes evident, however, when transmitter release from the HN is potentiated by intracellular administration of TEA or by cooling the preparation below 7°C , which reversibly suppresses the slow i.p.s.p. Figure 1c shows the i.p.s.ps before and 5 min after injection of TEA when the duration of the presynaptic action potential had more than doubled and the i.p.s.ps were substantially increased in duration and amplitude. The one-to-one nature of the fast i.p.s.p. and its constant latency at 4°C are illustrated in Fig. 1d and e, respectively. In other experiments where the slow i.p.s.p. was also eliminated by cooling, it was possible to reverse the polarity of the fast i.p.s.ps at membrane potentials greater than -65 mV (Eck) (ref. 4). Thus, the sensitivity to cooling of the slow i.p.s.ps, a known characteristic

of potassium-dependent potentials in *Aplysia*⁵, and the polarity change of the fast i.p.s.p. below -65 mV suggest that these fast and slow i.p.s.ps are mediated by chloride and potassium conductances, respectively.

Another class of postsynaptic response is illustrated in Fig. 2. Endogenous spiking activity in these follower neurones was inhibited by HN stimulation and by the application of histamine. Although these followers are functionally inhibited by spike activity in the HN, a small excitatory component always precedes the hyperpolarising response. This early component is more apparent when the membrane potential of the follower neurone is polarised to prevent spiking. The sample records shown in Fig. 2 b–d illustrate the postsynaptic responses at greater amplification and at different membrane potentials. When the membrane potential was increased, the excitatory component (arrows) became more prominent and at -80 mV the response was almost entirely excitatory. Similarly, the hyperpolarising response to exogenously applied histamine decreased and at -80 mV the histamine response was biphasic. Thus it seems that two types of histamine receptors may exist on the membranes of these follower neurones, one mediating a hyperpolarising response and the other associated with depolarisation. The biphasic nature of the p.s.ps is further illustrated in Fig. 2e–h when the membrane potential of the follower neurone was preset to -60 and -90 mV. The synaptic responses shown reveal several characteristics of these biphasic responses: (1) they are composed of a rapid (200–400 ms) depolarising component (e.p.s.p.) and a slowly developing and prolonged (s) i.p.s.p.; (2) the equilibrium potential of the i.p.s.p. is near -80 mV because it was reversed in polarity at -90 mV

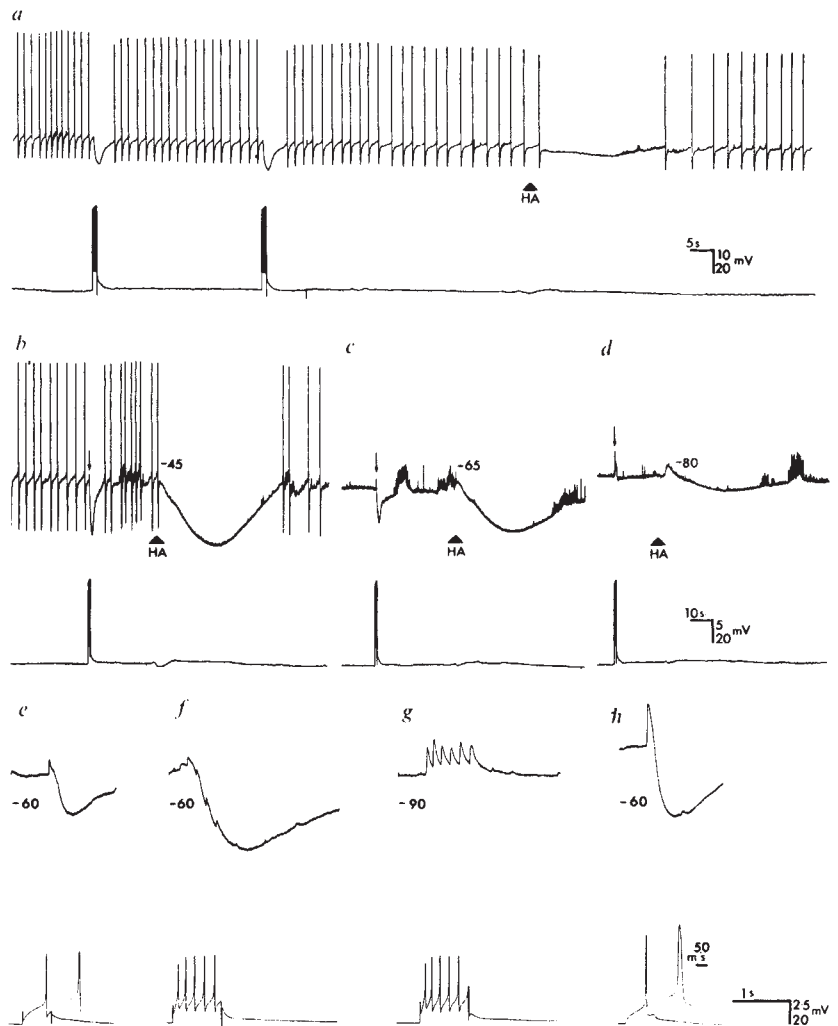


Fig. 2 Biphasic responses in another class of follower neurones. *a*, Effects of trains of HN spikes and the application of histamine (HA) on spontaneous spiking activity of the follower neurone recorded at resting potential (-45 mV); histamine applied by diffusion from a micropipette. *b–d*, Effects of HN spikes and histamine recorded at various membrane potentials shown in each set; arrows indicate the onset of the early excitatory synaptic component of the evoked response. *e–g*, Biphasic p.s.ps recorded at different membrane potentials and after various numbers of HN spikes. *h*, Biphasic p.s.p. 30 min after TEA was intracellularly injected into the HN; insets near lower traces show duration of HN spikes before (*e*) and 30 min after TEA administration (*h*). All records from the same follower neurone.

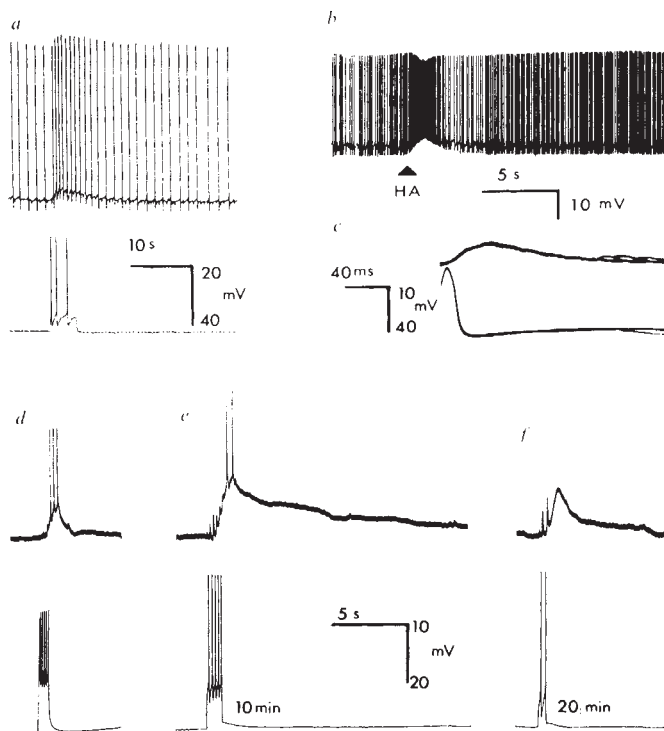


Fig. 3 Excitatory responses elicited by HN stimulation. *a*, Effects of three HN spikes on the endogenous activity of a follower neurone. *b*, Excitatory responses to iontophoretic application (20 nA, 0.5 s) of histamine (HA) delivered to the somal membrane of another follower neurone receiving excitatory synaptic input from HN. *c*, Four superimposed traces of e.p.s.ps and HN spikes recorded from the same follower as illustrated in (*b*). *d-f*, Effects produced by intracellular TEA into the HN in another preparation bathed in ASW containing normal (10 mM) Ca^{2+} ; membrane potential preset to -55 mV. *d*, Control responses; *e* and *f*, obtained at times indicated after administration of TEA. At 20 min after injection, two HN spikes produce fast and slow e.p.s.ps, with amplitude and duration substantially larger than control response to a train of seven HN spikes. E.p.s.ps in *d* and *e* cause firing of the follower neurone.

(Fig. 2g) and nearly absent at -80 mV (Fig. 2d); and (3) the amplitude of the i.p.s.ps, unlike those of the e.p.s.ps, were potentiated by repetitive firing of the HN (compare traces *f* and *g* in Fig. 2). Figure 2h shows the increased HN spike duration and the accompanying increases in the amplitudes and durations of both e.p.s.p. and i.p.s.p. components of the biphasic p.s.p. 30 min after TEA injection.

In a third class of follower neurones, stimulation of the HN produced excitatory responses, that is an increase in the rate of spiking (Fig. 3a); these followers also show an excitatory response to iontophoretically applied histamine (Fig. 3c). Like the two previously described synaptic responses, these excitatory followers also received two-component p.s.ps—a fast depolarising component (fast e.p.s.p.) several hundred milliseconds in duration and a long-lasting (seconds) depolarising component (slow e.p.s.p.). The sample records shown in Fig. 3c illustrate the constant latency between the HN spikes and the fast e.p.s.p. The two-component nature of these excitatory synaptic responses is clearly observed when transmitter release is enhanced by prolonging the HN spike by injection of TEA or by lowering the temperature of the bathing medium (data not shown). The responses shown in Fig. 3d-f reveal the effects of intracellular injection of TEA on the e.p.s.ps recorded from another preparation bathed in normal ASW (10 mM Ca^{2+}). Ten minutes after injection the amplitude and duration of both fast and slow e.p.s.ps were substantially increased and by 20 min the responses were even further potentiated.

These results demonstrate that single HN excite some follower neurones and inhibit others, and that the synaptic potentials underlying these effects are multicomponent, consisting of fast and slow p.s.ps in various combinations: fast i.p.s.p.+slow i.p.s.p.; fast e.p.s.p.+slow i.p.s.p.; and fast e.p.s.p.+slow e.p.s.p. Multiphasic p.s.ps elicited by single presynaptic neurones are not an uncommon feature of molluscan synapses⁶⁻¹¹, and multi-receptor systems for proven or suspected neurotransmitters have been characterised on molluscan neurones¹²⁻¹⁵. The tests used to distinguish between polysynaptic and monosynaptic connections indicate that all synaptic responses evoked by HN are mediated by monosynaptic pathways. Even though the measurements of latency and constancy of the slow synaptic responses are compromised by their gradual onset and long duration, the persistence of the slow p.s.ps. in high- Ca^{2+} ASW and their enhancement produced by intracellular injection of TEA is consistent with the monosynaptic nature of these slow responses.

The observations that neurones receiving monosynaptic connections from the HN respond to exogenously applied histamine in a manner qualitatively similar to the effects produced by HN stimulation demonstrates that histamine could act as a neurotransmitter at these identified synapses. This view is strengthened by the observations that histamine-sensitive neurones are but a small fraction of the neurones within this nervous system (ref. 16 and my unpublished results). Thus from the results described here together with previous biochemical data^{1,2} showing that these same presynaptic neurones uniquely synthesise and contain substantial amounts of histamine, it seems that this system of identifiable nerve cells will be useful for clarifying whether histamine is actually released as a synaptic neurotransmitter.

I thank the NSF for financial support and Dr N. Brookes for helpful discussions and criticism of the manuscript.

DANIEL WEINREICH

Department of Pharmacology and
Experimental Therapeutics,
University of Maryland,
School of Medicine,
660 West Redwood Street,
Baltimore, Maryland 21201

Received 2 March; accepted 2 May 1977.

- Weinreich, D., Weiner, C. & McCaman, R. E. *Brain Res.* **84**, 341-345 (1975).
- Weinreich, D. & Yu, Y.-T. *J. Neurochem.* **28**, 361-369 (1977).
- Berry, M. S. & Pentreath, V. W. *Brain Res.* **105**, 1-20 (1976).
- Russell, J. M. & Brown, A. M. *J. gen. Physiol.* **60**, 499-518 (1972).
- Kehoe, J. S. in *Drug Receptors* (ed. Rang, H.) 63-84 (University Park Press, Baltimore, 1973).
- Berry, M. S. & Cottrell, G. A. *J. Physiol.* **244**, 589-612 (1975).
- Gardner, D. & Kandel, E. R. *Science* **176**, 675-678 (1972).
- Kandel, E. R. & Gardner, D. *Res. Publ. Ass. nerv. ment. Dis.* **50**, 91-144 (1972).
- Kehoe, J. S. *Nature* **221**, 866-868 (1969).
- Shimahara, T. & Tauc, L. *J. Physiol., Lond.* **247**, 299-319 (1975).
- Wachtel, H. & Kandel, E. R. *Science* **158**, 1206-1208 (1967).
- Ascher, P. *J. Physiol.* **225**, 173-209 (1972).
- Gerschenfeld, H. M. & Paupardin-Tritsch, D. *J. Physiol., Lond.* **243**, 427-456 (1974).
- Kehoe, J. S. *J. Physiol.* **225**, 115-146 (1972).
- Swann, J. & Carpenter, D. O. *Nature* **258**, 751-754 (1975).
- Carpenter, D. O. & Gaubatz, G. L. *Nature* **254**, 343-344 (1972).

Reduced β -adrenergic receptor concentrations in ageing man

CONCENTRATIONS of certain hormone receptors are known to be reduced during ageing in several animal species¹⁻⁴. Most relevant work, however, has been limited to intracellular hormone receptors such as those for steroids, and less is known about changes in cell surface receptors, including those for catecholamines. We have measured the concentration of β -adrenergic receptors in crude membrane fractions of mononuclear cells from human subjects of various ages. Receptor affinities for ($-$)³H-dihydroalprenolol were essentially the same for all subjects. But a

significant decrease in receptor concentration was observed in cells from older individuals.

Subjects were 23 males between 24 and 81 yr old, without significant illnesses and taking no medications. Some were paid volunteers and others were recruited from the longitudinal study of ageing of the National Institute on Ageing and thus were living in a community and fully ambulatory. The study lasted 6 months with a random distribution of subjects by age. We made multiple determinations of receptor concentrations on several subjects during 2–3 months. Subjects fasted and did not smoke for 12 h before the study. With written consent, 112–225 ml of blood was withdrawn by antecubital venipuncture into heparinised syringes, with subjects in the supine position. Whole blood was subjected to density gradient centrifugation by the method of Böyum, using a Ficoll–dextrazote mixture (specific gravity 1.077)⁵. Mononuclear cell fractions thus obtained were assayed for β -adrenergic receptors by a modification of the method of Williams *et al.*⁶. This method uses the labelled β -adrenergic antagonist (–)³H-dihydroalprenolol incubated at various concentrations (10^{-9} to 10^{-7} M) with washed crude membrane preparations in the presence and absence of 10^{-5} M unlabelled (–) dihydroalprenolol to determine specific binding. Numbers of lymphocytes were determined by morphological criteria using Turks solution and the nonspecific esterase staining technique adapted from Yam *et al.*⁷. Cell viability was evaluated by Trypan blue exclusion. Protein determinations were made by the method of Lowry *et al.*⁸.

The specific binding of (–)³H-dihydroalprenolol to the washed crude membrane preparation was saturable, of high affinity and stereospecific in favour of the (–) stereoisomer. These findings are consistent with those of Williams *et al.*, who showed that binding detected by this method possesses the characteristics of the β -adrenergic receptor⁶. Lymphocytes constituted 83% of the mononuclear cells obtained. Viability as assessed by Trypan blue exclusion was generally greater than 94%. Cell yield, cell viability and protein yield per cell did not vary with the age of the donor. Saturation plots of β -adrenergic receptor concentrations in crude mononuclear membrane preparations of two representative subjects are shown in Fig. 1. At saturation levels (greater than 5×10^{-8} M (–)³H-dihydroalprenolol) specific binding accounted for approximately 50% of total binding, although at lower concentrations the percentage of specific binding was greater, as reported by Williams *et al.*⁶.

The receptor has high affinity for (–)³H-dihydroalprenolol in preparations from both young and old subjects, with K_{DS} of approximately 20 nM for each group. The mean saturation level of young subjects (24–41 yr, $n = 11$), however, was 572 ± 58 pmol mg⁻¹ protein ($\pm$ s.e.m.) while for subjects older than 46 yr ($n = 12$) the corresponding value was slightly greater than half that value,

332 ± 48 pmol mg⁻¹ protein ($\pm$ s.e.m. $P < 0.01$). When all data are considered (Fig. 2), there is a significant negative correlation of receptor concentration with age ($r = 0.641$, $P < 0.001$).

Receptor number per cell also correlated inversely with age from approximately 14,000 to 8,000 sites per cell in the young and old groups, respectively. Receptor levels in individual subjects remained fairly constant over short periods (2–3 months).

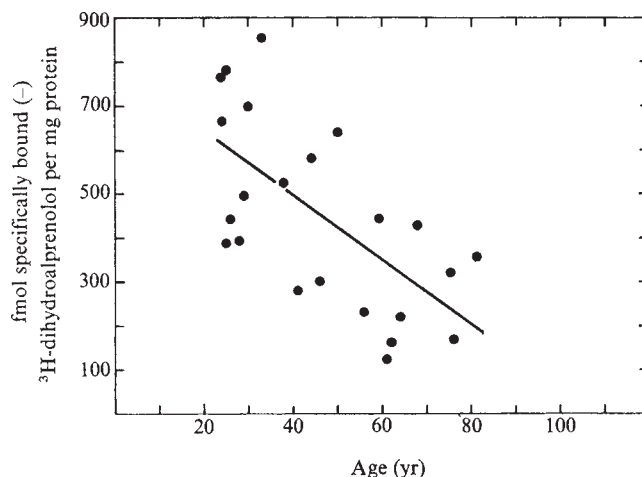


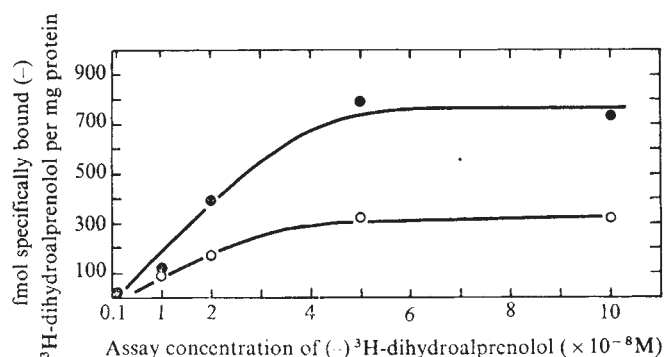
Fig. 2 Maximal specific binding of (–)³H-dihydroalprenolol to crude mononuclear cell membranes of subjects aged 24–81 yr.

The search for physiological and biochemical determinants of ageing has led to investigations of age-associated alterations in hormone receptors. Steroid receptor concentrations are known to decline with age in various rodent tissues^{1,2,4,9,11}, and the ability of steroid hormones to elicit physiological responses has been shown to be diminished with age in several of these tissues^{1,10,11}. Whereas steroid hormone receptors are intracellular, some receptors located on the cell surface also seem to change with age. Concentrations of rat hepatic insulin receptors and adipocyte glucagon receptors, however, seem to decline very early in life^{12,13}.

Previous work with human material has dealt with either intracellular receptors or cells in tissue culture. Binding of glucocorticoids to intracellular receptors in human liver obtained postmortem is diminished in older individuals³. The effect of age on binding of insulin to cultured human fibroblasts has also been studied, but the results have been controversial^{14,15}. Our investigation is the first to demonstrate age-associated alterations of surface hormone receptors in cells taken directly from man.

The aetiology of these receptor alterations remains unclear. Differences in the composition of mononuclear cell populations or individual changes in the mononuclear cells might result in variation between young and old individuals. Relative and absolute numbers of T and B lymphocytes have been thought to remain fairly stable during human ageing^{16,17}, but this has been questioned⁴. Selective removal of monocytes from the mononuclear cell preparation by a nylon mesh column does not markedly influence the amount of (–)³H-dihydroalprenolol specifically bound⁶. The values obtained for the concentration of β -adrenergic receptors of young (age 24–41 yr) subjects, while higher than those initially published⁶, compares more closely with values currently obtained using similar methods (ranging from 96 to 476 fmol (–)³H-dihydroalprenolol specifically bound per mg protein) (personal communication from K. Kariman and R. J. Lefkowitz). Additionally, diminished concentrations of β -adrenergic receptors have been reported in frog erythrocyte membranes previously

Fig. 1 Specific binding of (–)³H-dihydroalprenolol to crude mononuclear cell membranes of representative subjects. ●, 24-yr-old; ○, 75-yr-old.



exposed to catecholamines *in vivo*¹⁹. This catecholamine-induced 'desensitisation' is another possible aetiology for receptor decline in ageing man, where plasma catecholamine levels are known to increase²⁰.

Altered responsiveness to catecholamines has been observed during ageing in several rodent and human organ systems as well as in intact man²¹⁻²⁴. Several immunological functions of lymphocytes are known to be influenced by adrenergic agents²⁵. The possible role of age-associated receptor reduction or 'desensitisation' in altering the effects of catecholamines on these functions requires elucidation.

Our findings complement previous studies in several aspects. Our results lend further support to the idea that loss of certain hormone receptors is a frequent manifestation of ageing. The demonstration of diminished β -adrenergic receptor concentrations in ageing man offers a possible explanation for the observed age-associated loss of adrenergic responsiveness.

D. D. SCHOCKEN
G. S. ROTH

Gerontology Research Center,
National Institute on Aging,
Baltimore City Hospitals,
Baltimore, Maryland 21224

Received 28 February; accepted 26 April 1977.

- ¹ Roth, G. S. & Livingston, J. N. *Endocrinology* **99**, 831-839 (1976).
- ² Shain, S. A., Boesel, R. W. & Axelrod, L. R. *Archs Biochem. Biophys.* **167**, 247-263 (1973).
- ³ Singer, S., Ito, H. & Litwack, G. *Int. J. Biochem.* **4**, 569-573 (1973).
- ⁴ Nelson, J. F. *et al. Abstr. 29th A. Meeting Gerontol. Soc.* **34** (1976).
- ⁵ Böyum, A. *Scand. J. clin. Invest.* **21**, Suppl. 97, 77-89 (1968).
- ⁶ Williams, L. T., Snyderman, R. & Lefkowitz, R. J. *J. clin. Invest.* **57**, 149-155 (1976).
- ⁷ Yam, L. T., Li, C. Y. & Crosby, W. H. *Am. J. clin. Path.* **55**, 283-290 (1971).
- ⁸ Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).
- ⁹ Kanungo, M. S., Patnaik, S. K. & Koul, O. *Nature* **253**, 366-367 (1975).
- ¹⁰ Moudgil, V. K. & Kanungo, M. S. *Biochem. biophys. Res. Commun.* **52**, 725-730 (1973).
- ¹¹ Roth, G. S. *Biochim. biophys. Acta* **399**, 145-156 (1975).
- ¹² Freeman, C., Karoly, K. & Adelman, R. C. *Biochem. biophys. Res. Commun.* **54**, 1573-1580 (1973).
- ¹³ Livingston, J. N., Cuatrecasas, P. & Lockwood, D. H. *J. Lipid Res.* **15**, 26-32 (1974).
- ¹⁴ Rosenbloom, A., Goldstein, S. & Yip, C. *Abstr. 5th int. Cong. Endocr., Hamburg*, 261 (1976).
- ¹⁵ Itoh, H., Orimo, H. & Shimada, H. *Abstr. 5th int. Cong. Endocr., Hamburg*, 261 (1976).
- ¹⁶ Augener, W., Cohnen, G., Reuter, A. & Brittinger, E. *Lancet* **i**, 1164 (1974).
- ¹⁷ Weksler, M. E. & Hutteroth, T. H. *J. clin. Invest.* **53**, 99-104 (1974).
- ¹⁸ Diaz-Jouanen, E., Williams, R. C. Jr & Strickland, R. G. *Lancet* **i**, 688-689 (1975).
- ¹⁹ Mukherjee, C., Caron, M. G. & Lefkowitz, R. J. *Proc. natn. Acad. Sci.* **72**, 1945-1949 (1975).
- ²⁰ Ziegler, M. G., Lake, C. R. & Kopin, I. J. *Nature* **261**, 333-335 (1976).
- ²¹ Fleisch, J. H., Maling, H. M. & Brodie, B. B. *Circulat. Res.* **26**, 151-162 (1970).
- ²² Berger, M., Preiss, H., Hesse-Wortmann, C. & Gries, F. *Gerontologia* **17**, 312-322 (1971).
- ²³ Lakatta, E. G., Gerstenblith, G., Angell, C. S., Shock, N. W. & Weisfeldt, M. L. *Circulat. Res.* **36**, 262-269 (1975).
- ²⁴ Yin, F. C. *et al. Circulation* **54**, Suppl. II, 167 (1976).
- ²⁵ Bourne, H. R. *et al. Science* **184**, 19-28 (1974).

Tuning properties of cochlear hair cells

THE mammalian cochlea is generally considered to be a frequency analyser. When the ear is stimulated by a pure tone the basilar membrane vibrates. The amplitude of the vibrations varies along the length of the membrane with the position of the maximum depending in a graded way upon the frequency of the tone^{1,2}. The basal turn responds maximally to high frequencies, the apex to low frequencies. Consequently hair cells situated along the membrane are frequency selective. The nerve fibres innervating the basilar membrane have low thresholds over a narrow range of frequencies corresponding to the displacement maxima of the section of the basilar membrane that they innervate³. But the neural tuning curves (the relationship between threshold and frequency for individual nerve fibres) are much sharper than the mechanical tuning curves of the basilar membrane (the relationship between amplitude of displacement and frequency at a single point on the membrane) and this discrepancy has led to the proposal of a 'second filter'⁴. In this paper we describe the tuning of intracellularly recorded receptor potentials from histologically identified hair cells (IHC) of the guinea pig cochlea

in an effort to investigate the nature of the hypothetical sharpening process and we conclude that IHCs are tuned as sharply as auditory nerve fibres. To our knowledge this is the first time intracellular recordings have been made from inner hair cells in the mammalian cochlea.

Guinea pigs weighing 120-200 g were maintained throughout the experiment under pentobarbitone-phenoperidine anaesthesia and were paralysed with Flaxedil (gallamine triethiodide) and artificially ventilated (method of E. F. Evans, personal communication). The cochlea was exposed ventrolaterally through the acoustic bulla. A fine Teflon insulated silver wire was positioned close to the round window and secured to the skull with cyanoacrylic cement so that the threshold (within 10 dB) of the compound action potential (AP) evoked by pure tone bursts of 2 ms duration could be monitored during the exposure of the basilar membrane. The sound source was a driven half-inch condenser microphone (B&K 4133) in a hollow ear bar which was inserted in the tympanic ring and secured to the temporal bone and adjacent bones with cyanoacrylic and dental cement. The basilar membrane of the basal turn was exposed taking care to minimise bleeding into the scala tympani thus preserving AP threshold. All results presented in this paper are uncorrected for the sound system which is measured to be flat within 10 dB. The sound pressure levels (SPL) indicated are referred to 0 dB SPL re 2×10^{-5} N m⁻².

Microelectrodes were drawn from 1 mm o.d. fibre filled boro-silicate glass tubing on a Livingstone type electrode puller. They had resistance of about 700 M Ω when filled with 6% procion yellow, but this was reduced to around 300-400 M Ω by bevelling the electrode tips (on a grinding wheel) to facilitate penetration of the basilar membrane⁵. Despite the reduction in their resistance by bevelling, the frequency response of the electrodes was too low to record cochlear microphonics and we have therefore concentrated on measuring the d.c. receptor potential (summing potential). The electrodes were advanced under visual control into the basilar membrane with a hydraulic micro-

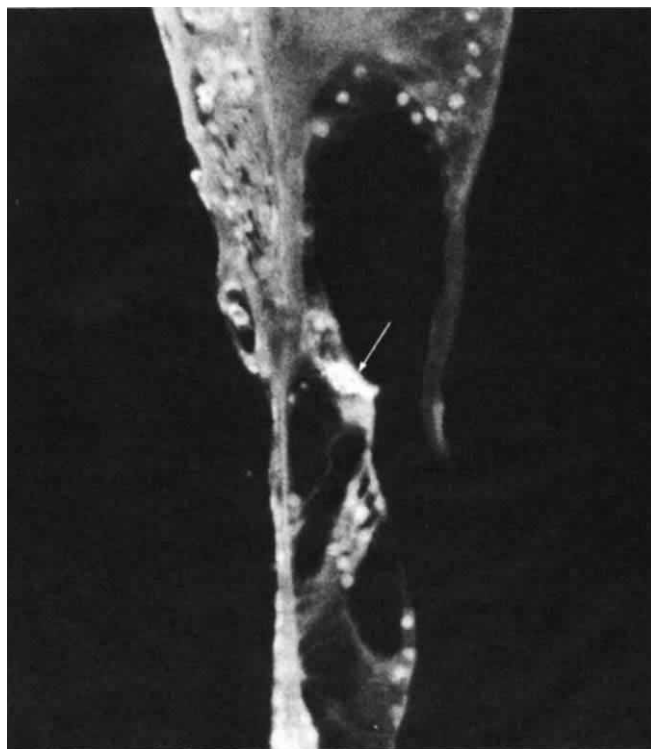


Fig. 1 An inner hair cell (arrow) filled with Procion Yellow in the basal turn of the right cochlea of animal 128. The photomicrograph was taken with ultraviolet illumination and the scale bar indicates 25 μ m.

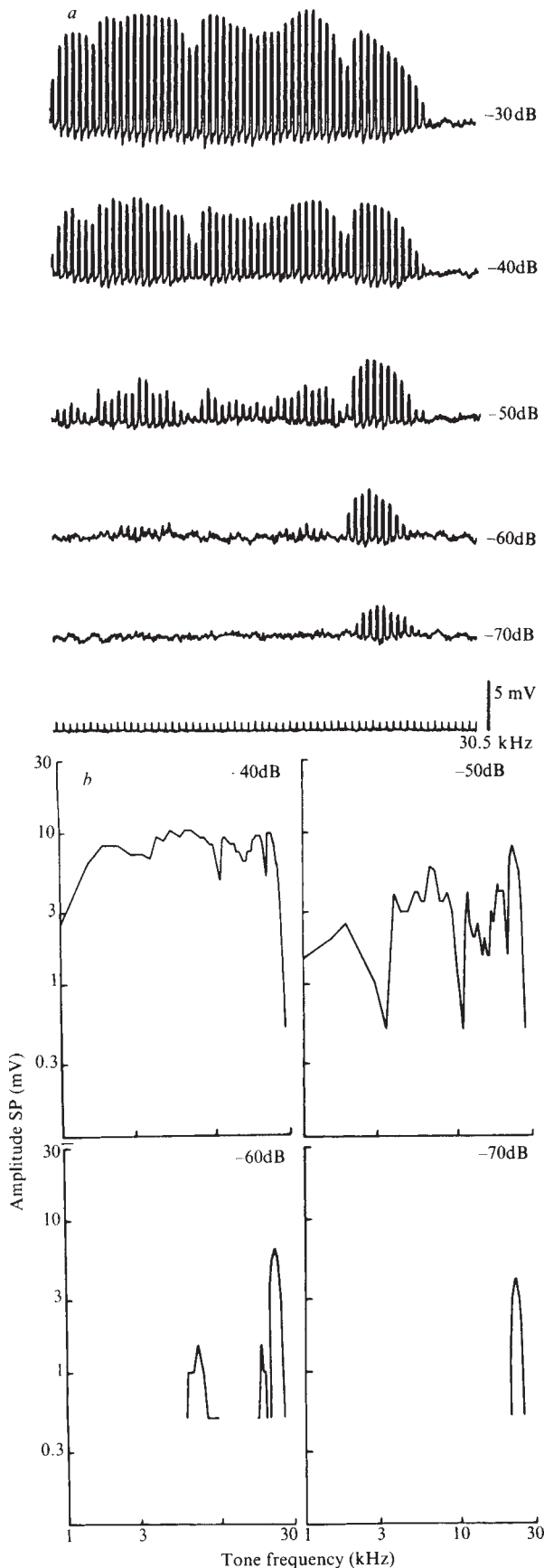


Fig. 2 *a*, Receptor potentials recorded intracellularly from an identified inner hair cell in the basal turn of the right cochlea of animal 105 in response to 15-ms tone bursts (rise and fall time 2 ms) presented at 470-Hz intervals. Trigger pulses for the tone bursts are shown in the lower trace beginning at 1 kHz and ending at 30.5 kHz. *b*, Isointensity curves of the receptor potentials shown in (*a*) plotted at their different SPL.

drive and were accurately positioned when the perilymph was momentarily drained. On rare occasions there was a transient rise in threshold after draining the scala tympani, however sensitivity always returned as the scala tympani refilled.⁶

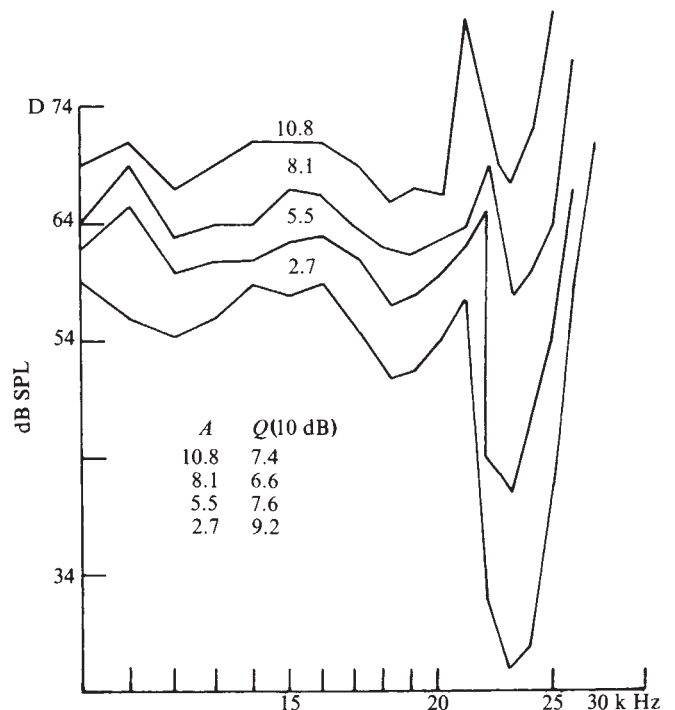
The first two or three cells encountered by the micro-electrode in the region of the IHCs had resting potentials of 80–90 mV and small depolarising potentials of 0–1 mV at 74 dB SPL at characteristic frequency (CF). When the electrode was advanced further, cells were encountered with resting potentials between 35–45 mV and with positive receptor potentials between 5 and 17 mV at 74 dB SPL at CF; negative receptor potentials were never recorded intracellularly.

Membrane potentials were sometimes stable for 10 min or more, but on occasion dropped within a minute of penetration. The magnitude of the receptor potential seemed to vary with the amount of damage sustained by HCs during penetration and was not highly correlated with the AP threshold. Of 26 dye marked cells with large receptor potentials, nine were successfully found under the microscope and were identified as IHCs. An example is shown in Fig. 1. Only the data from these cells are discussed here.

The frequency response of the IHCs was determined between 1 and 40 kHz by presenting 15-ms tone bursts at 470-Hz intervals. At high SPL (74 dB), the frequency response is similar to that of the basilar membrane. However, as the SPL is reduced the receptor potentials at frequencies below the CF fall faster than those at the CF so that at 44 dB SPL the cell is very sharply tuned. Figure 2 illustrates the response of a cell with a CF of 23 kHz.

The interpretation of our results rests largely on the relative importance of the d.c. receptor potential (summing potential) and cochlear microphonic in exciting the afferent nerve fibres. At AP threshold the receptor potential can be as large as 1.6 mV and increases with the level of stimulation to a maximum of about 17 mV, 70 dB above threshold. These recorded potentials are larger than receptor potentials recorded in hair cells in other organs^{7,8} and compare in magnitude with presynaptic potentials recorded elsewhere in the nervous systems⁹. Thus, in view

Fig. 3 Isoamplitude curves of intracellular receptor potentials from experiment 105. *A*, Amplitude of receptor potential in mV.



of their size, it may be that the d.c. receptor potential is responsible for exciting the auditory nerve. If this is so, then the tuning properties of the auditory nerve depend upon the tuning properties of the receptor potential.

To determine whether the frequency response of the intracellular receptor potential corresponds to that of auditory nerve fibres, the receptor potential data must be converted into isoamplitude curves since neural tuning curves are plots of threshold against frequency (isorate). Isoamplitude plots of receptor potential against frequency converted into isoamplitude curves since neural tuning curves are level dependent and, therefore, a comparison with neural tuning curves must be made at a receptor potential that corresponds to the neural threshold. Since AP threshold (which corresponds to single fibre threshold) was 10 dB below the lowest SPL at which receptor potentials were measured, the lowest isoamplitude plot (2.7 mV) is considered to be the closest to the neural threshold tuning curve.

The sharpness of auditory nerve tuning curves is measured in terms of the Q_{10dB} (CF divided by the bandwidth 10 dB above the threshold at CF) and this reaches a maximum of about 10 in the most sensitive fibres³. The Q_{10dB} of the 2.7 mV isoamplitude curve is 9.2 and we therefore conclude that intracellularly recorded receptor potential is tuned as sharply as the auditory nerve and hence that any sharpening of the frequency response beyond that of the basilar membrane mechanics does not involve neural interaction.

We thank Professor R. J. Andrew and Dr K. Boshier for their encouragement and loan of equipment. We thank Drs T. S. Collett and M. F. Land for reading the manuscript and Miss Wendy Randall for technical assistance. This work was supported by the MRC and by an equipment grant from the charities panel of the Readers' Digest.

I. J. RUSSELL
P. M. SELICK

*Ethology and Neurophysiology Group,
School of Biology,
University of Sussex, Falmer, UK*

Received 24 January; accepted 29 April 1977.

¹ Bekesy, G. von *Experiments in Hearing* (McGraw-Hill, New York, 1964).

² Johnstone, B. M. & Boyle, A. J. F. *Science* **158**, 389–390 (1967).

³ Robertson, D. & Manley, G. A. J. *comp. Physiol.* **91**, 363–375 (1974).

⁴ Evans, E. F. J. *Physiol., Lond.* **226**, 263–287 (1972).

⁵ Brown, K. T. & Flaming, D. D. *Brain Res.* **86**, 172–180 (1975).

⁶ Robertson, D. *Science* **186**, 135–155 (1974).

⁷ Harris, G. G., Frischkopf, L. S. & Flock, A. *Science* **167**, 76–79 (1970).

⁸ Mulroy, M. J., Altmann, D. W., Weiss, T. F. & Peake, W. T. *Nature* **249**, 482–485 (1974).

⁹ Eccles, J. C. *The Physiology of Synapses* (Academic, New York, 1964).

Lysogeny by f2 phage?

Two prominent properties of lysogenic bacteria are, first, inheritable potentiality for phage production and, second, immunity against homologous phage. Both characteristics are unaffected by growth in the presence of phage anti-serum or by serial subcultures of single colonies. This indicates that every lysogenic cell carries the genetic information for the biosynthesis of a given type of the phage particle¹; so far, however, this has been known only for some DNA bacteriophages. I present here the first demonstration that bacteria can also contain the genetic information for the production of an RNA phage. If *Escherichia coli* bacteria are grown and infected with f2 RNA phage in liquid medium supplemented with guanidinium chloride, many cells subsequently behave as true lysogens. They are 'immune' to f2 and spontaneously release phage identical in morphology, plaque type, serological property, male specificity and single-stranded RNA genome, to the phage used for the initial infection.

f2 is a very small virulent phage and is male specific—

adsorbing only to F pili of male *E. coli* strains. Its single-stranded RNA chromosome has two functions: it is the infectious part of the phage particle, and acts as viral messenger in the infected cell that controls the synthesis of two structural proteins and RNA replicase². This dual function for viral single-stranded RNA is also found in animal poliovirus. Guanidine prevents the synthesis of viral RNA by guanidine-sensitive strains of poliovirus by inhibiting completely the replication of viral genome³. Whether guanidinium chloride has a similar effect in bacteria, that is, inhibition of replication of f2 RNA, is not known.

Guanidinium chloride at a final concentration of 10 mM in Tryptone liquid medium⁴, does not interfere with growth and colony forming ability of *E. coli* bacteria. But when bacteria grown in the presence of guanidinium chloride were infected by f2 in otherwise optimal conditions for f2 growth⁴ the yield of phage was strongly reduced and the percentage of bacteria surviving was markedly increased (Table 1).

To distinguish between resistance and immunity, the guanidinium chloride-treated and f2 infected (multiplicity of infection < 1) male bacteria were allowed to grow with aeration for 3 h after the infection and plated on G-agar (agar⁴ supplemented with 10 mM guanidinium chloride) with anti-phage serum. The colonies thus obtained were picked out, and, when plated on G-agar with the indicator bacteria in the top agar⁴, some of them (about 16%) gave rise to plaques. It is possible that the serum did not inactivate all the phage attached to the F pili. To check this, aliquots of f2-producing bacteria were inoculated into liquid G-medium supplemented with anti-phage serum (at the concentration that inactivated more than 99% of the

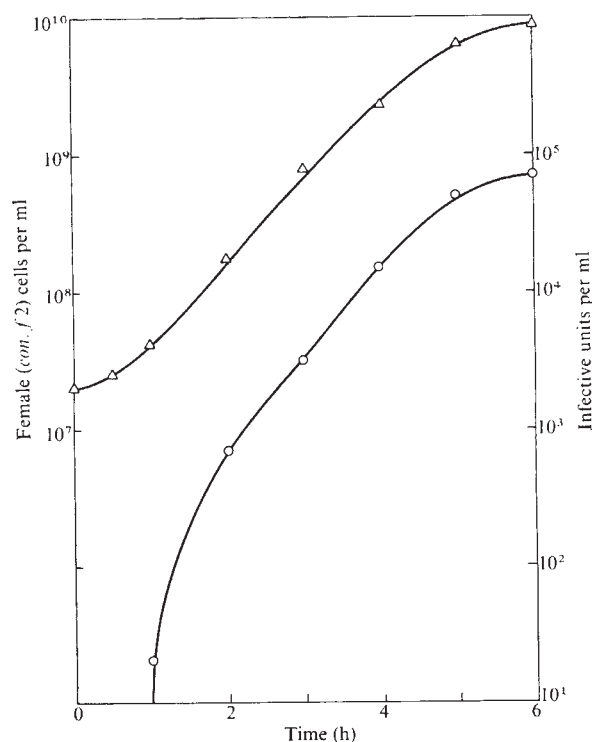


Fig. 1 Spontaneous release of the phage by *E. coli*(con.f2) cells. F⁺(con.f2) bacteria were grown overnight in G-medium supplemented with anti-phage serum at the concentration that inactivates more than 99% phage during 5 min. After two washings by centrifugation in saline-calcium and treatment with phage antiserum⁴, the bacteria were inoculated into fresh G-medium (zero time) and incubated with aeration up to the stationary phase of growth. The number of viable bacteria (Δ) and of infective units (○) were estimated hourly (as described in Table 1). Essentially the same results were obtained with male(con.f2) bacteria.

Table 1 Immunity of (*con. f2*) bacteria to f2 phage

Bacterial culture	³ H-leucine-labelled phage ³ H- sedimented with		³ H-uridine-labelled phage ³ H- sedimented with		f2 yield per ml of infected culture	Bacteria surviving infection (%)
	Antiserum (%)	Bacteria (%)	Antiserum (%)	Bacteria (%)		
Hfr RA-2	86	5	14	77	7.8×10^{11}	26
Hfr RA-2, G-medium	82	8	11	75	6.1×10^9	84
Hfr RA-2(<i>con. f2</i>)	84	7	13	78	1.3×10^6	91

To label phage RNA, ³H-uridine (specific activity 7.4 Ci mmol⁻¹, Radiochemical Centre, Amersham) was added to bacterial culture 30 min before infection, at 2.5 µCi ml⁻¹; to label phage coat, ³H-leucine (150 mCi mmol⁻¹, Radiochemical Centre) was added to bacterial culture 30 min before infection, at 2.5 µCi ml⁻¹. Phage stocks were prepared¹ and purified by density gradient centrifugation in CsCl. The phage labelled in its RNA or protein were used separately, at a multiplicity of 0.8, to infect *E. coli* Hfr RA-2 (from Dr K. B. Low), growing either in Tryptone medium or in Tryptone medium supplemented with 10 mM guanidinium chloride (G-medium), and *E. coli* Hfr RA-2(*con. f2*) growing in G-medium at 37 °C (using experimental procedure for f2 growth¹). 10 min after the infection, free phage (that is, phage still attached to F pili, phage ghosts and truly free phage) were removed by violent agitation in a Towers shaker for 3 min. Thereafter the infected bacteria were freed from free phage by two washings of the cells by centrifugation in saline-calcium² and sedimented by centrifugation at 3000 r.p.m. for 5 min (Sorval centrifuge, rotor SS34). Free phage (in the supernatant) were treated with f2 antiserum for 10 min at 37 °C and sedimented at 10,000 r.p.m. for 30 min. The pellet of infected cells was resuspended into fresh medium (Tryptone or G-medium) and that of free phage into saline-calcium. One aliquot of the bacterial suspension was used for the estimation of radioactivity³ (in Table 1: ³H sedimented with bacteria), and other for the examination of bacterial survival⁴ and phage production by one cycle of growth¹, using Hfr RA-2 as an indicator strain. Free phage were used for the estimation of radioactivity³ (in Table 1: ³H sedimented with antiserum). The results represent average values from three separate experiments. Essentially the same results were obtained with other male bacteria: *E. coli* F⁺CR 63 from Dr M. Radman, and *E. coli* F⁺ts lac⁺ pro⁺/lac⁻ pro⁻ from Dr D. J. Cummings (the latter grown at 30 °C). f2 phage was donated by Dr A. D. Kaiser.

phage during 5 min), grown to stationary phase and then plated again with antiserum. After five successive re-isolations (after long growth with anti-phage serum in each case) bacteria showing completely new characteristics were isolated: (1) they were not resistant to f2, but did not produce the phage when infected, as though they had been 'immune' to f2 (Table 1) and (2), during the growth in the liquid medium they released phage (Fig. 1). These bacteria are designated (*con. f2*), that is, containing the genetic information for f2 biosynthesis.

It is unlikely that anti-phage serum treatment did not remove all the phage and that some phage still remained

adsorbed to F pili of male bacteria. If this were to occur, however, it would not explain the 'immunity' observed. Definite proof that the phenomenon described here is not due to f2 particles adsorbed to F pili was obtained by constructing F⁻(*con. f2*) bacteria using F⁺ts (temperature sensitive) cells for infection. At the inhibitory temperature, the thermosensitive F⁺ plasmid is diluted out from the *E. coli* population, because its replication is inhibited and it is inherited unilinearly by the descendants of the original F⁺ bacteria⁵. One aliquot of *E. coli* F⁺ts lac⁺ pro⁺/lac⁻ pro⁻ cells was allowed to grow in G-medium at 42 °C, while the other was infected in the same medium at 30 °C with f2, and after 10 min warmed to 42 °C by adding fresh medium prewarmed at 60 °C. The growth of cells was continued at 42 °C for 3 h, and infected and non-infected bacteria were plated at 42 °C on MacConkey agar—enabling us to distinguish male from female bacteria⁶. Colonies of female cells were picked out and F⁻(*con. f2*) bacteria were made as above for (*con. f2*) male cells; the only difference was that bacteria were grown at 42 °C. Controls on MacConkey agar were used at each step of F⁻ and F⁻(*con. f2*) strain isolation. The female cells thus obtained were lac⁻ pro⁻, and they could not adsorb male-specific phage f2, R17 and MS2. F⁻(*con. f2*) bacteria, however, produce phage spontaneously (Fig. 1). This phage is, by its morphology (Fig. 2), male specificity, serological property, single-stranded RNA genome (examined by the method of Bradley⁷) and plaque type (on Tryptone agar), identical to the phage f2. But on G-Tryptone agar mottled plaques occur, from the turbid zone of which male (*con. f2*) bacteria can be isolated. The functional state and the location of genetic information for phage production in (*con. f2*) bacteria, and the molecular mechanism of guanidinium chloride action are under investigation.

I thank Drs A. D. Kaiser and D. J. Cummings for discussions, Dr Durda Novak for antiserum preparation, Dr Mercedes Wrischer for electron micrographs, and Miss Josipa Hrženjak for technical assistance.

VERA ZGAGA

Laboratory of Cellular Radiobiology,
Institute 'Rudjer Boskovic',
41001 Zagreb, Yugoslavia

Received 29 March; accepted 4 May 1977.

- Jacob, F. in *Les bacteries lysogènes et la notion de provirus* 21–22 (Masson et Cie, Paris, 1954).
- Lodish, H. F., Cooper, S. & Zinder, N. D. *Virology* 24, 60–70 (1964).
- Summers, D. F., Maizel, J. V. & Darnell, J. E. *Proc. natn. Acad. Sci. U.S.A.* 54, 505–513 (1965).

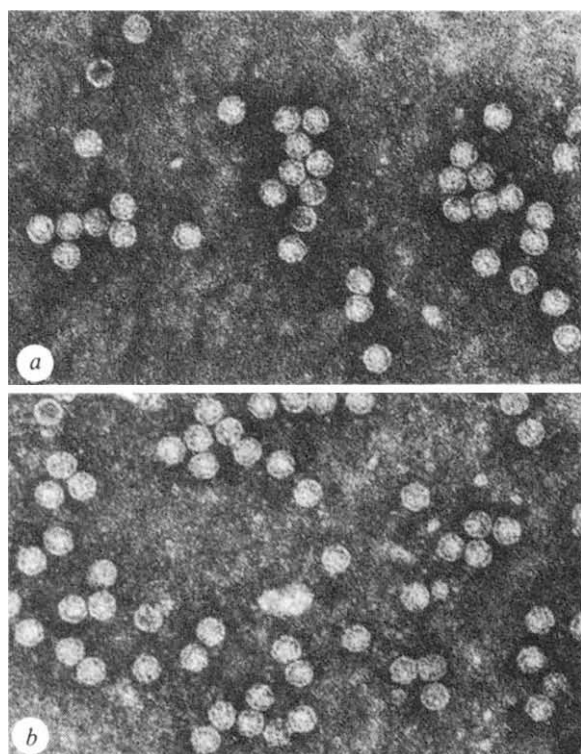


Fig. 2 Electron micrographs showing the morphology of phage. Phage, prepared by negative staining, were examined in a Siemens Elmiscop 1. a, Wild type f2 particle; b, phage released spontaneously by F⁻(*con. f2*) cells. (×150,000.)

- 4 Cooper, S. & Zinder, N. D. *Virology* **18**, 405–411 (1962).
 5 Cuzin, F. & Jacob, F. *Ann. Inst. Pasteur* **112**, 397–418 (1967).
 6 Braun, W. in *Bacterial Genetics* 100–111 (W. B. Saunders, London, 1953).
 7 Bradley, D. E. *J. gen. Microbiol.* **44**, 383–391 (1966).
 8 Zgaga, V., Medić, M., Salaj-Smic, E., Novak, D. & Wrischer, M. *J. molec. Biol.* **79**, 697–708 (1973).
 9 Adams, M. H. in *Bacteriophages* (Wiley-Interscience, New York, 1959).

Ether induced morphological alteration of Pf-1 filamentous phage

BACTERIOPHAGES of filamentous structure have been isolated in several species of bacteria and all have similar physico-chemical properties¹. They consist of single-stranded circular DNA and two species of proteins—A protein and B protein. They are released from infected cells without disintegrating the cells and are inactivated on exposure to chloroform although the presence of lipids in the virion has not been reported. During a study of Pf-1, a filamentous phage which propagated on strain K of *Pseudomonas aeruginosa*², we found that it is almost completely inactivated with ether, whereas phage fd and Xf are refractory to it. Ether sensitivity of the viruses is generally thought to be an indication of the presence of lipids in the virion. Pf-1 might, therefore, be thought to contain lipids in its structure. We report here a study of the inactivation and morphological alteration of Pf-1 induced by ether and the results of chemical analysis of Pf-1 which suggest the absence of lipids in this virus.

For the propagation of Pf-1 and fd, nutrient broth was used and Xf was propagated in the medium described by Kuo *et al.*³. The effects of organic solvents on the plaque forming activity of filamentous phages are shown in Table 1. Pf-1 was completely inactivated by ether and other solvents tested, whereas fd and Xf were only inactivated with chloroform.

Table 1 Inactivation of filamentous phages by organic solvents

Treatment	Titre (PFU ml ⁻¹)		
	Pf-1	Xf	fd
None	1.3×10^{11}	1.6×10^{11}	4.2×10^{11}
Ethyl ether	$< 10^3$	1.5×10^{11}	3.6×10^{11}
Chloroform	$< 10^3$	8.0×10^7	1.5×10^7
Methanol	$< 10^3$	NT*	4.2×10^{11}
Acetone	$< 10^3$	NT	3.0×10^{11}

Culture fluid containing about 10^{11} PFU ml⁻¹ phage was mixed with an equal volume of solvent. The mixture was stirred gently with a pipette and allowed to stand for 10 min at room temperature. Phage activity in the aqueous phase of the mixture separated by low-speed centrifugation (500g, 5 min) was titrated using a plaque technique.

*NT, Not tested.

Electron microscopic observation of ether treated Pf-1 showed that phage filaments disappeared and many rod-like structure appeared (Fig. 1a). This morphological change observed in Pf-1 could be the reason for the loss of the activity of this phage. The width of the rod was 185 Å and the length was variable but always shorter than that of the phage filament. The middle portion of the rod was densely stained with phosphotungstate. From these electron microscopic images two types of structure for the rod can be postulated. One is a hairpin like structure which can be formed by joining two phage filaments bent like a hairpin, and the other is a hollow cylinder formed by the swelling of a phage filament. Shadowing techniques revealed the surfaces of the rod-like structures, and the grooves which would be expected from the images of hairpin like structures were not observed (Fig. 1b). The effects of ether on the activity of Pf-1 were rapid, and after exposure to the ether vapour for even a few seconds a decrease of its activity could be measured. Electron microscopic observation of this preparation showed that phage filaments were swelling into rod-like structures at a portion of a

filament. Penetration of the stain occurred on some of them, giving the appearance of the rod-like structures seen in Fig. 1a (Fig. 1c). These results suggest that ether treatment causes swelling of the phage into a hollow cylindrical structure.

To investigate the chemical composition of Pf-1, phage was purified from infected culture fluid with the method of ammonium sulphate fractionation and equilibrium density gradient centrifugation of CsCl. Bacteria were removed from 2 l of fluid by centrifugation (8,500g for 30 min), 240 g of ammonium sulphate was added gradually with stirring to precipitate the phage.

The precipitate was collected by centrifugation (8,500g, 30 min) and dialysed against 0.01 M Tris-HCl buffer pH 7.4 overnight. The cell debris that remained in the phage

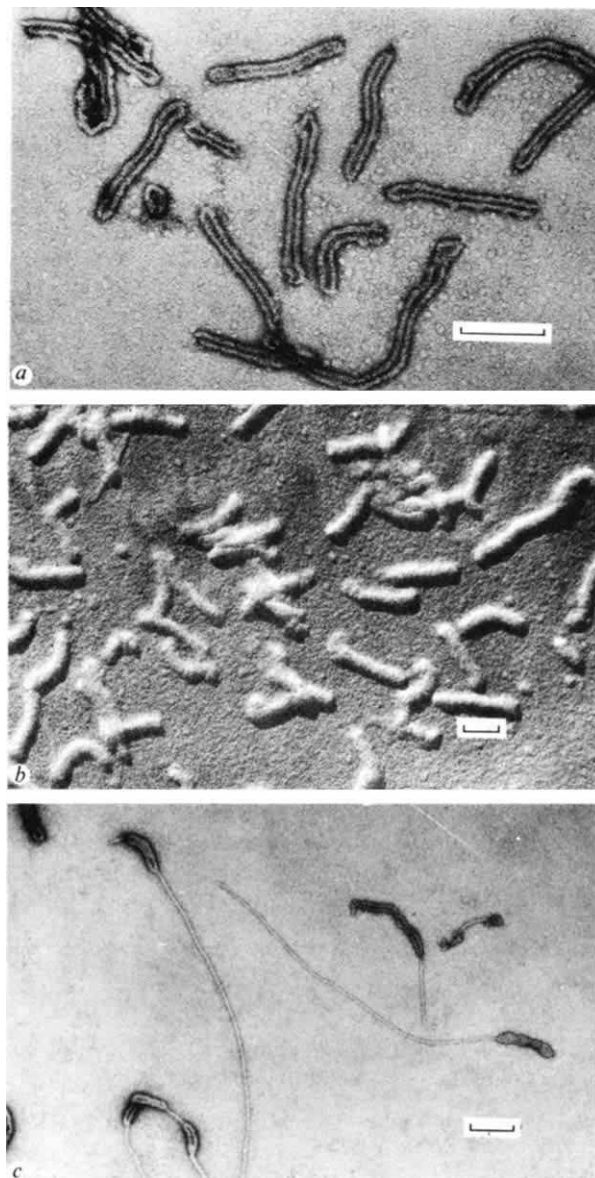


Fig. 1 Electron micrographs of Pf-1 treated with ether. *a*, Negatively stained Pf-1 treated with ether. The purified phage was mixed with an equal volume of ether for 10 min at room temperature. The aqueous phase was negatively stained with potassium phosphotungstate, pH 7.2, on the electron microscopic grid prepared as described by Fukami and Adachi⁸. Scale bar represents 100 nm. *b*, The same preparation of Pf-1 as in (*a*) was shadowed with tungstate at an angle of 30° using an Edwards E306 vacuum evaporator equipped with an electron beam device. *c*, Negatively stained Pf-1 treated with ether vapour. Purified phage was placed on specimen grids and exposed to ether vapour for 5 s by keeping the grids above the test tube containing ether.

suspension was removed by three cycles of differential centrifugation (20,000g, 10 min and 100,000g, 2 h). The gradient consisting of three steps (1.4, 1.3 and 1.2) and spun at 84,000g for 4 h. The phage band which appeared in the lower third of the gradient was collected, dialysed against 0.01 M Tris-HCl buffer overnight and further purified by CsCl equilibrium sedimentation (115,000g, 21 h). The phage band at 1.29 density was collected and dialysed against the buffer and this is referred to as the purified phage preparation. This phage preparation contained 1.6×10^{13} PFU per ml (plaque forming units) of phage which was 20% of the activity detected in the infected culture fluid. To demonstrate the purity of the phage preparation it was subjected to polyacrylamide gel electrophoresis containing sodium dodecyl sulphate (SDS) to separate the proteins following the method of Dunker and Rueckert⁴.

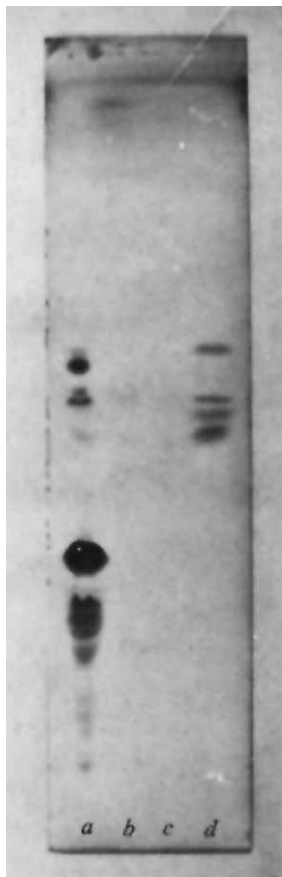


Fig. 2 Thin-layer chromatography of the lipids fraction extracted from Pf-1. Lipids were extracted from the purified phage suspension or bacterial cells by the method of Braunstein and Franklin⁵. Purified phage (1×10^{13} PFU) or bacterial cells (2×10^{10} cells) were suspended in 4.8 ml of 0.01 M Tris-HCl buffer, pH 7.4. Methanol-chloroform (2:1, v/v, 18 ml) was added to the suspension and intermittently shaken for 2 h at room temperature. The precipitated residue was centrifuged at 3,000g for 20 min and the supernatant retained. Extractions were repeated in the residue. Chloroform (6 ml) was added to the supernatant and mixed. Then 6 ml of water was added and also mixed. The emulsion was centrifuged at 3,000g for 30 min and the chloroform layer was collected, dried in room temperature and retained as the lipid fraction. Lipid extracts dissolved in the mixture of chloroform and methanol were applied 2 cm from one edge of a chromatographic plate. The plate was placed in a cylindrical jar containing 10 ml of chloroform-methanol-water (65:25:4, v/v) and allowed to stand until the front of the solvent reached 11 cm from the edge of the plate. The plate was air-dried and then transferred to the other jar containing 10 ml of the mixture of *n*-hexane-ethyl ether-acetic acid (90:10:1, v/v) and developed further until the solvent front reached 19 cm from the edge of the plate. The spots of lipid were visualised by staining with iodine or molibden blue. *a*, Lipid extracts from strain k of *Ps. aeruginosa*; *b*, lipid extracts from Pf-1; *c*, solvent alone; *d*, standard lipids, from top to bottom; triolein, diolein, cholesterol and oleic acid.

Only the two visible bands of protein which represented A protein and B protein respectively were stained. No other bands indicating the contamination of the proteins from bacteria were detected.

Lipid was extracted from the purified Pf-1 by the method of Braunstein and Franklin⁵ using 2.5 mg of phage (1×10^{13} PFU). Lipid fractions were analysed on a thin layer chromatography plate coated with silica gel 60 (Merck). Using the same plate the lipid fractions extracted from host bacteria (2×10^{10} cell) were analysed simultaneously. The plate was initially developed with a solvent mixture consisting of chloroform-methanol-water (65:25:4, v/v) and then with *n*-hexane-ethyl ether-acetic acid (90:10:1, v/v). The plate was stained with iodine or molibden blue; several spots were visualised on the specimens extracted from bacteria but none on the extracts from Pf-1. This implied that Pf-1, although it is sensitive to ether, does not contain significant amounts of lipid in its structure. Significant levels of lipids have been found in many other ether-sensitive viruses.

In lipid-containing viruses the morphological alterations induced by ether-treatment are thought to be caused by the extraction of the lipids which have a role in maintaining the morphological integrity of the virion. In Pf-1, the morphological alteration seems not to be the disintegration of the phage filament but its swelling into a hollow cylinder.

The structure of filamentous phages including Pf-1 has been extensively examined by X-ray diffraction and chemical analysis. A structural model has been presented by Marvin⁶.

The protein shell of the filamentous phage is constructed from coat protein molecules exhibiting largely an α -helix structure and extending in an axial direction with the phage filaments overlapping each other like fish scales to form a hollow cylindrical tube, in which the virus DNA is wrapped. In the phage protein shell, the coat protein molecules were held together by the hydrophobic bond created by the hydrophobic region of the coat protein molecules⁷. It is assumed that ether probably disrupts these bonds and induces deformation of the arrangement of the coat proteins on the protein shell and results in the swelling of phage filament. Our preliminary experiment to observe the DNA molecules in ether-treated phage specimens by the technique of protein monolayer method revealed the presence of many small circular DNA molecules similar in size and shape to those extracted by the phenol method from the purified Pf-1. This suggests that the DNA molecules packed in the protein shell might be released concomitantly with the swelling of the phage filament.

KAZUNOBU AMAKO
KAKUKO YASUNAKA

Department of Microbiology,
School of Medicine,
Fukuoka University, Nishi-Ku,
Fukuoka 814, Japan

Received 25 April; accepted 11 May 1977.

¹ Marvin, D. A. & Hohn, B. *Bact. Rev.* 33, 172-209 (1969).

² Takeya, K. & Amako, K. *Virology* 28, 163-165 (1966).

³ Kuo, T. T., Hutang, T. C., Rong, Y. W. & Yang, C. M. *Botan. Bull. Acad. Sinica* 8, 246-254 (1967).

⁴ Dunker, A. K. & Rueckert, R. R. *J. biol. Chem.* 244, 5074-5080 (1969).

⁵ Braunstein, S. N. & Franklin, R. M. *Virology* 43, 685-695 (1971).

⁶ Marvin, D. A. & Wachtel, E. J. *Nature* 253, 19-23 (1975).

⁷ Nakashima, Y., Weiseman, R. L., Konigsberg, W. & Marvin, D. A. *Nature* 253, 68-71 (1975).

⁸ Fukami, A. & Adachi, K. *J. Electron Microscopy* 14, 112-118 (1965).

Aflatoxin B₁-oxide generated by chemical or enzymic oxidation of aflatoxin B₁ causes guanine substitution in nucleic acids

AFLATOXIN B₁ (AFB₁) has the highest biological activity of the four naturally occurring aflatoxins produced by the

mould *Aspergillus flavus*¹. Feeding studies have shown it to be the most potent liver carcinogen known for the rat². Its ingestion by man is associated with the high primary liver cancer incidence in certain parts of Africa³. There is much evidence to suggest that AFB₁ requires metabolism to exert both its carcinogenic and mutagenic effects (reviewed in ref. 4). Structure-activity⁵, as well as chemical studies^{6,7} indicate that the major route for activation proceeds through mixed function oxidase attack yielding the 8,9-oxide (previously called the 2,3-oxide but renumbered according to IUPAC recommendations). Attempts to chemically synthesise this metabolite have so far been unsuccessful. We report here that peracid oxidation of AFB₁ generates AFB₁-8,9-oxide, and that this latter compound reacts readily with nucleic acids to give adducts identical to those obtained after liver mixed function oxidase (MFO) activation *in vitro* and *in vivo*.

Using a modification of a chemical procedure recommended for the epoxidation of acid sensitive olefins⁸, we are able to covalently bind AFB₁ to DNA, polyguanylic acid or deoxyguanosine. Enzymic binding to nucleic acid was achieved by incubating the carcinogen with rat or hamster liver post-mitochondrial supernatants and DNA together with the necessary cofactors.

Figure 1 shows the chromatographic profile on Sephadex LH-20 of deoxynucleosides obtained by enzymic hydrolysis of DNA after MFO activation of AFB₁ compared with that after chemical activation. Four peaks of radioactivity were obtained, in each case the major peak being product I. Freeze-drying of fractions from peaks I and II and injection onto a silica reversed-phase column and subsequent high pressure liquid chromatography showed each peak to be

homogeneous (data to be presented elsewhere).

Mild acid hydrolysis of AFB₁-reacted DNA, after MFO activation, to release purine bases and subsequent LH-20 chromatography of the hydrolysate, showed three peaks of radioactivity. Two of the peaks had similar retention times to the minor peaks (III and IV) obtained after enzymic hydrolysis. A similar chromatographic profile after mild acid hydrolysis was obtained for AFB₁-reacted DNA after chemical activation (data not shown).

Evidence for the structures of the various products of AFB₁ reaction with DNA was obtained by substituting deoxyguanosine or polyguanylic acid as nucleophiles, in place of DNA, in the chemical oxidation system. Chromatography of the deoxyguanosine reaction products on Sephadex LH-20 gave peaks of radioactivity coincident with peaks I, II and IV obtained in enzymic hydrolysis of DNA after either chemical or enzymic reaction with AFB₁ (Fig. 2). Incubation of the deoxyguanosine reaction products at pH 9.0 for 24 h and subsequent Sephadex LH-20 chromatography showed peaks II and IV to largely disappear and be replaced by peaks I and III. Strong acid hydrolysis (1 N HCl, for 60 min at 100 °C) of AFB₁-reacted polyguanylic acid from a chemical activation experiment and Sephadex LH-20 chromatography of the reaction products showed two major peaks of radioactivity (III and V) and a minor one (IV).

Thus, all the reaction products obtained after neutral enzymic hydrolysis of DNA, except for V, seem to be derived from guanine. Peak V is that of AFB₁-8,9-dihydrodiol, since it is excluded at the fifth void volume on Sephadex LH-20 chromatography⁹, and co-chromatographs with synthetically prepared diol. Acid hydrolysis of DNA

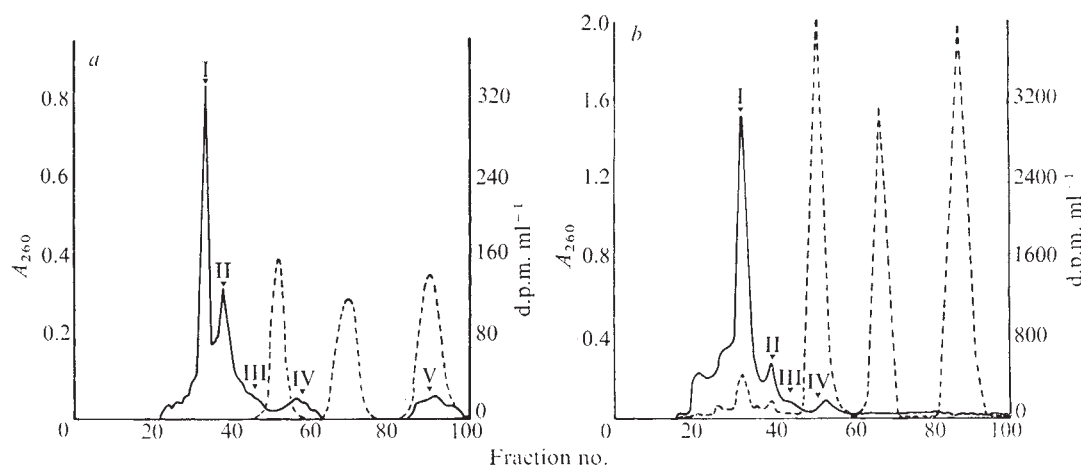


Fig. 1 Chromatographic profiles of neutral enzymic digests of DNA on Sephadex LH-20. —, ¹⁴C radioactivity; ---, A₂₆₀. *a*, Enzymic activation of AFB₁. Hamster or rat liver post-mitochondrial supernatant (1 ml of a 25% homogenate in 150 mM KCl, equivalent to 250 mg of fresh liver) was incubated in 100 mM sodium phosphate buffer pH 7.4 containing 1.5 μmol NADP⁺, 20 μmol glucose-6-phosphate, 25 μmol MgCl₂, 10 mg of calf thymus DNA and 200 μg ¹⁴C-AFB₁ dissolved in 0.2 ml dimethyl sulphoxide (specific radioactivity 0.52 mCi mmol⁻¹) in a total volume of 3 ml. After incubation at 37 °C for 60 min, the incubation mixture was extracted once with an equal volume of phenol-cresol, the nucleic acid precipitated from the aqueous layer with ethoxyethanol and redissolved in 5 ml sodium chloride-sodium acetate buffer. Contaminating RNA was removed by incubation with ribonuclease-A (50 μg ml⁻¹) and the DNA re-extracted with 0.5 volumes phenol-cresol. The nucleic acid was then precipitated, redissolved and chloroform extracted three times before centrifugation at 100,000g for 60 min to remove glycogen. 1.7 mg AFB₁-reacted DNA dissolved in 1 ml 0.3 M sodium chloride-0.25 M sodium acetate-0.015 M magnesium acetate was digested with 100 μg deoxyribonuclease I for 60 min at 37 °C. After the addition of 3 units bacterial alkaline phosphatase and 0.05 units snake venom phosphodiesterase the mixture was incubated for a further 18 h at pH 7.0 and 37 °C. *b*, Chemical oxidation of AFB₁. 10 mg of calf thymus DNA were dissolved in 8 ml 20 mM sodium phosphate buffer pH 7.4 and added to a solution containing ¹⁴C-AFB₁, dissolved in 8 ml dichloromethane (67.5 μg ml⁻¹ ¹⁴C-AFB₁, specific radioactivity 0.4 mCi mmol⁻¹ prepared as previously described¹⁴). To this two-phase system was added 625 μg 3-chloroperbenzoic acid (4 molar excess) and the whole stirred for 25 h at room temperature in the dark in a stoppered flask. At the end of this time, the two phases were separated, and the aqueous phase, containing the nucleic acid, extracted five times with equal volumes of chloroform. The nucleic acid was then precipitated in three volumes of ethanol and redissolved in 20 mM Tris HCl pH 7.0. This ethanol precipitation step was repeated three times and the nucleic acid finally redissolved in 0.3 M sodium chloride/0.25 M sodium acetate. 4 mg of the AFB₁-reacted DNA dissolved in 1 ml sodium chloride-sodium acetate-magnesium acetate were digested with twice the enzyme concentrations in (*a*). Enzyme digests were chromatographed on Sephadex LH-20 columns (110 × 1.75 cm) eluted with water, 4-ml fractions collected and an aliquot counted for radioactivity. Absorbance was measured at 260 nm. The main ultraviolet absorbing peaks from left to right are deoxycytosine and thymidine combined, deoxyguanosine and deoxyadenosine. The results are from two similar columns run separately.

or poly G to which AFB₁ has been bound, leads to release of this derivative, showing that AFB₁-8,9-oxide is formed by either enzymic or chemical oxidation of AFB₁.

3-Chloroperbenzoic acid is a highly specific oxidising agent for the epoxidation of olefinic double bonds. Use of this reagent to epoxidise AFB₁ in the presence of DNA, results in a high percentage of nucleic acid bases being modified. No evidence was obtained for direct action of the peracid on DNA. In the conditions outlined here, one molecule of AFB₁ is bound per 100 molecules of base, thus, approximately one in every 20 guanine bases are modified. Increasing the amount of AFB₁ with a proportionate increase in the concentration of peracids leads to a higher level of binding. With liver MFO, lower amounts of binding were found but this was again dependent on the AFB₁ concentration.

Neutral enzymic hydrolysis of DNA leads to the release of deoxynucleoside adducts⁹, whereas mild acid hydrolysis results in the release of base adducts¹⁰. Peaks I and II seem therefore to be deoxyguanosine adducts of AFB₁-8,9-oxide whereas III and IV are guanine adducts of the epoxide. Mild acid hydrolysis (0.1 N HCl at 70 °C for 20 min) of peak I converts it to peak III (data not shown) suggesting

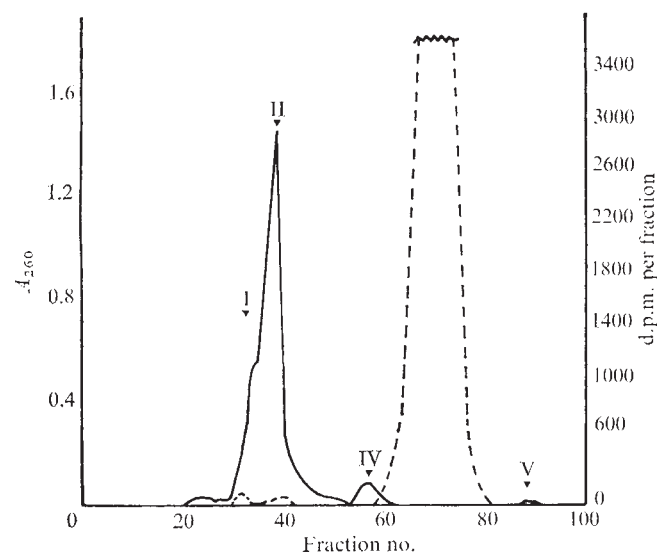


Fig. 2 Chromatographic profile of deoxyguanosine AFB₁-reacted adducts after chemical oxidation of AFB₁ at pH 7.4. —, ¹⁴C radioactivity; - - -, A₂₆₀. Deoxyguanosine was reacted with AFB₁ as in Fig. 1b, with the exception of the ethanol precipitation step. Chromatography on Sephadex LH-20 was as in Fig. 1.

that peak III results from the loss of a sugar residue from peak I. Incubation of peak II at 37 °C, pH 9.0 converts it to peak I while peak IV is converted to III. Two explanations for this are that (1) alkali attacks the base-AFB₁ adduct to open the lactone ring, or (2) alkali attacks the adduct to open the imidazole ring of guanine. Explanation (1) is unlikely since peak I has the characteristic ultraviolet absorption maximum of AFB₁ between 360 and 370 nm and also incubation of AFB₁ at pH 9.0 does not result in lactone ring opening. Explanation (2) is substantiated by the absorption spectrum of peak I at pH 7.0 in the range 260–280 nm which is characteristic of a pyrimidine, suggesting destruction of the imidazole ring of the purine moiety.

Imidazole ring opening of guanine in mild alkaline conditions is characteristic of N-7 substituted guanines¹¹. Hydrolysis of AFB₁-reacted-DNA at neutrality for 30 min at 100 °C, results in the removal of the AFB₁-guanine adduct. Chromatography of the heat-hydrolysed supernatant after acid precipitation of the pyrimidine oligonucleotides at 0 °C on Sephadex LH-20 shows peaks III and

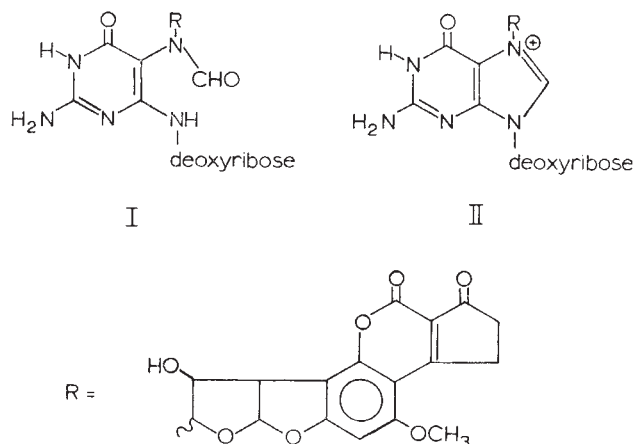


Fig. 3 Postulated structures of products I and II.

IV have been released (data not shown). Again this is a characteristic of guanine alkylated at the N-7 position¹².

Reaction at other positions in guanine is ruled out by (1) the relative ease of imidazole ring opening, (2) depurination at 100 °C in neutral conditions, (3) the ease of release of AFB₁-8,9-dihydrodiol from the nucleoside-AFB₁ adducts in mild acid conditions and (4) the base-AFB₁ adducts are stable at 100 °C in 1 N HCl. These findings lead us to assign the structures in Fig. 3 to the various products formed on reaction of AFB₁-8,9-oxide with DNA. Peaks I and II are the deoxyguanosine adducts whereas peaks III and IV are the guanine adducts. Data to be presented elsewhere shows, in agreement with other workers that products I to V are also formed on hydrolysis of liver DNA after ¹⁴C-aflatoxin B₁ administration to rats. Work is in progress to obtain sufficient of the products for chemical characterisation as well as identifying any possible minor reaction products.

The results reported here indicate that AFB₁-8,9-oxide is formed during peracid oxidation or MFO activation of AFB₁ and that this metabolite reacts with nucleic acid. Two major products of reaction are found, as has also been shown by Swenson *et al.*¹³, and these are both derived from guanine, almost certainly through N-7 substitution.

Covalent binding of AFB₁ with guanine at the N-7 position is likely to induce frame-shift mutations due to intercalation of the AFB₁ molecule together with base-substitution mutations as a result of misrepair. This reaction may therefore be important in the initiation of cancer as well as mutation induction in microorganisms by this potent mycotoxin.

This work was supported by the MRC and the Yorkshire Cancer Research Campaign. We thank Drs J. R. Lindsay Smith and B. F. Coles for advice.

CARL N. MARTIN
R. COLIN GARNER

Cancer Research Unit,
University of York,
Heslington, Yorkshire, UK

Received 25 February; accepted 29 March 1977.

- Wogan, G. N. in *Methods in Cancer Research* (ed. Busch, H.) 309–344 (Academic, New York, 1973).
- Wogan, G. N., Pagliarunga, S. & Newberne, P. M. *Fd. Cosmet. Toxicol.* **12**, 681–685 (1974).
- Peers, F. G., Gilman, G. A. & Linsell, C. A. *Int. J. Cancer* **17**, 167–176 (1976).
- Campbell, T. C. & Hayes, J. R. *Toxicol. appl. Pharm.* **35**, 199–222 (1976).
- Garner, R. C., Miller, E. C. & Miller, J. A. *Cancer Res.* **32**, 2058–2066 (1972).
- Swenson, D. H., Miller, E. C. & Miller, J. A. *Biochem. biophys. Res. Commun.* **60**, 1036–1043 (1974).
- Garner, R. C. *FEBS Lett.* **36**, 261–264 (1973).
- Anderson, W. K. & Veysoglu, T. J. *Org. Chem.* **38**, 2267–2268 (1973).
- Miller, C. T., Lawley, P. D. & Shah, S. A. *Biochem. J.* **136**, 387–393 (1973).
- Lawley, P. D. & Thatcher, C. A. *Biochem. J.* **116**, 693–707 (1970).
- Lawley, P. D. & Brookes, P. *Biochem. J.* **89**, 127–138 (1963).
- Lawley, P. D. & Shah, S. A. *Chem.-Biol. Interact.* **7**, 115–120 (1973).
- Swenson, D. H., Lin, J. K., Miller, E. C. & Miller, J. A. *Cancer Res.* **37**, 172–181 (1977).
- Garner, R. C. *Chem.-Biol. Interact.* **6**, 125–129 (1973).

matters arising

Repression of synaptic efficacy in frog skeletal muscle

GRINNELL, Rheuben and Letinsky¹ show convincingly that when the transplanted frog sartorius muscle is innervated by two foreign nerves, control of the muscle fibres may be shared between the two, one nerve being effective at transmission the other eliciting only subthreshold junction potentials of low quantal content. They suggest this depends on a mutual competition between nerves similar to that invoked to explain selective reinnervation of urodele muscle^{2,3}.

I wonder whether this pattern of innervation is simply a property of the frog sartorius muscle. Luff and Proske⁴ have found using tension measurements, that there is very little overlap between motor units when they measure twitch tension but extensive overlap (50% or more) when tetani were used. Furthermore, motor units in which the twitch tension fluctuated from trial to trial required high rates of stimulation to achieve peak tetanic tension and had low twitch tetanus ratios. This suggests that normally many muscle fibres are supplied with junctions which have a low safety factor for neuromuscular transmission. Presumably those fibres are also supplied with a safe junction from another motoneurone. Since there are only about 12 motor axons supplying the sartorius it seems unlikely, although not impossible, that the reduced efficacy of some of the junctions depends on an embryological matching process such as that postulated to account for selective innervation of muscles after nerve regeneration in urodeles. Rather it may be because sartorius muscle fibres tend to maintain one neuromuscular junction at high efficacy and reduce the efficacy of others on the same fibre, regardless of the embryological origin of the synapses, an interesting developmental mechanism in its own right but not necessarily related to the formation of specific connections. One could answer the question by finding out whether in the preparation used by Grinnell *et al.* there were unequal connections made by different motoneurons within the two populations of foreign nerve fibres. The test would be to split one of the foreign nerves and see if there was evidence for competition between different populations of fibres in one nerve, as there is between

the populations of fibres of the two foreign nerves.

RICHARD MARK

*Department of Behavioural Biology,
Research School of Biological Sciences,
P.O. Box 475,
Canberra City ACT. 2601, Australia.*

¹ Grinnell, A. D., Rheuben, Mary B. & Letinsky, M. S. *Nature* 265, 368–370 (1976).

² Yip, J. W. & Dennis, M. J. *Nature* 260, 350–352 (1976).

³ Genat, B. R. & Mark, R. F. *Phil. Trans. R. Soc. B* 278, 335–346 (1977).

⁴ Luff, A. R. & Proske, U. *J. Physiol., Lond.* 258, 673–685 (1976).

GRINNELL *et al.* REPLY—Mark rightly asks the question: If competition between two different nerves reduces the effectiveness of many terminals, as we report, then should not the same phenomenon occur between axons of the same nerve? If so, might this not be a normal phenomenon whereby each fibre acquires (relatively) exclusive innervation by one axon? This was an obvious question to us as well, and we have done some experiments of the type he suggests.

Our results indicate that there is competition between axons of the same nerve, but much less than between axons of two different foreign nerves. When a single nerve (normal or reinnervated) is split into two or more segments, with one to three functional motor units in each, there is little if any twitch overlap (usually 5% or less), but up to 25% tetanus overlap. This implies the existence, in *Rana catesbeiana* as in *Litoria aurea*¹, of some subthreshold polyneuronal innervation, and leads us to agree with Mark that the competition we have demonstrated between pairs of foreign nerves may be a special case of similar competition between axons of a single nerve. In our experiments there is a sharp quantitative difference between competition in the two cases, however, and probably a qualitative difference. The tension overlap we find between axons in singly innervated muscles does not approach that reported by Luff and Proske¹. The explanation for this difference may be one of species or technique. It is noteworthy, for example, that Luff and Proske's Ringer apparently contained only 1.08 mM Ca²⁺ compared with the 1.8 mM in our amphibian Ringer and that used by most other investigators. In four experiments in our laboratory, shifting a normal sartorius preparation from 1.8 mM Ca²⁺ to 1.08 mM Ca²⁺ Ringer caused no change in 50 per cent tetanus tension, but a reversible 53 ±

28% drop in twitch tension, suggesting that quantal content in a high proportion of junctions can be reduced below threshold by that reduction in Ca²⁺ concentration. This could also help account for Luff and Proske's low twitch-tetanus ratios. Although we find a few subthreshold junctions in singly reinnervated muscles, normally the quantal content is close to threshold, never as low as is commonly found in doubly innervated muscles (under 10). Moreover, the average twitch tension developed by simultaneous stimulation of both of two foreign nerves is conspicuously less (in percentage of fibres twitching) than that elicited by stimulation of the nerve in singly reinnervated muscles. This implies either fewer fibres are innervated, or that there is a greater number of subthreshold junctions where two different nerves are present rather than only one. Since we normally find that either nerve when tetanised can drive almost 100% of the muscle, most of the junctions formed by each must be subthreshold².

The competition we describe between the foreign somatic motor nerves is nonspecific; neither has an apparent innate advantage. Thus we do not propose this as a direct explanation for the reported instances of specific repression of one nerve by another. On the other hand, we do see some distinction between axons of two different nerves, compared with those of a single nerve. Moreover, the competition occurs in a muscle with focal rather than distributed endings, and hence is amenable to direct study of single 'repressed' endings. It is our hope that an understanding of this interaction may provide important insights into more specific forms of interaction.

ALAN D. GRINNELL

*Department of Biology and
Brain Research Institute,
University of California, Los Angeles,
Los Angeles, California 90024*

MARY B. RHEUBEN

*Department of Biology,
Pennsylvania State University,
University Park, Pennsylvania 16802*

MICHAEL S. LETINSKY

*Department of Physiology and
Brain Research Institute,
University of California, Los Angeles,
Los Angeles, California 90024*

¹ Luff, A. R. & Proske, U. *J. Physiol., Lond.* 258, 673–684 (1976).

² Grinnell, A. D., Rheuben, Mary B. & Letinsky, M. S. *Nature* 265, 368–370 (1977).

reviews

Jumping into black holes

Roman Znajek

The Iron Sun: Crossing the Universe Through Black Holes. By Adrian Berry. Pp. 176. (Jonathan Cape: London, 1977.) £3.95.

ADRIAN BERRY has written a book about making black holes and jumping into them. He claims that the latter is a way of achieving instantaneous interstellar transport. After reading half the book one eventually discovers that it is called *The Iron Sun* because Berry proposes to build the holes by using magnetic fields to sweep together bits of iron (and other junk) that float around between stars. As the holes are to be ten times as massive as the Sun, and are to be situated at the 'safe' distance of one light year, this is clearly going to be a colossal task. Berry, however, thinks that humanity will continue to increase its economic output at an average annual rate of 3 per cent, and will therefore be able to achieve this feat "within about 250 years' time". Opposition is to be expected from assorted politicians, conservationists and second-rate scientists, but Berry is nevertheless confident that the project will be successfully completed by a conspiracy of dedicated technocrats.

I should explain that a black hole is a region of space where gravity is so strong that nothing, not even light, can escape. Its formation is inevitable when a sufficiently compressed piece of matter gets too cold to support itself against its own gravitational field. A ten solar mass lump of iron will certainly start to collapse under its own weight. What Berry does not seem to realise is that it will also get very hot. It is possible that the outer regions will stop collapsing and then get blown away. In other words, Berry may have created a recipe for an artificial supernova. A supernova one light year away would be a thoroughly unpleasant object and the conservationists would be justified in opposing it. There is of course no point in building the hole much further away because it is precisely in order to travel across such distances that the hole is being built in the first place.

An optimist may say that this is a mere technical difficulty and that there might be some way of cooling the matter as it collapsed. So how is a black hole going to help us get to the

stars? According to Berry, whenever a black hole is formed, a white hole 'simultaneously' appears in another part of the Universe. (A white hole is the opposite of a black hole: instead of falling into it, matter tends to fall out.) The two holes are supposed to be connected by a short piece of space that is 'outside' normal space—a kind of cosmic Suez Canal. A properly navigated spaceship can disappear into the black hole and reappear an instant later coming out of the white hole, even though a beam of light would take many years over the same journey in ordinary space.

Berry admits that many problems must be solved before such a voyage is possible. He confesses at the end of the book that no one knows where the white hole is going to appear. The fact is that no one knows *why* the white hole is going to appear. The idea that black holes are linked to white holes is a speculation which very few relativists have taken seriously. Speculation is legitimate, but quoting it as established knowledge is not. Berry has performed a grave disservice to science by not passing on warnings like: "concepts outside our present knowledge of physics must be invoked" which regu-

larly appear in the references he relies on. He owns up to the instability of the short cuts through which his spaceships travel, but does not draw the appropriate conclusion that the short cuts simply don't exist.

There are very many errors in this book. Perhaps the worst is Berry's claim that absolute time exists in the theory of special relativity. He says that although "many people who should know better" deny it, Einstein himself believed in absolute time and the simultaneity of events. This is an important issue for Berry, because if absolute time did not exist his instantaneous form of transport would lead to logical paradoxes. Unfortunately, the page to which Berry refers in Einstein's *The Meaning of Relativity* appears to be about something else, and if Berry looks at the preceding page he will no doubt decide that Einstein too needs to know better . . .

Adrian Berry should have known better than to write this book. □

Roman Znajek is an SRC Post-Doctoral Fellow at the Institute of Astronomy, University of Cambridge, and a Research Fellow of King's College, University of Cambridge, UK.

Immunopathology

Textbook of Immunopathology. Second Edition. Vol. 1: Pp. xlviii+1-554. Vol. 2: Pp. xlviii+555-1,118. Edited by P. A. Miescher and H. J. Muller-Eberhard. (Grune and Stratton: New York and London, 1976.) \$49.50; £30.20 each volume.

THIS textbook of immunological mechanisms in disease and immunopathology has become a standard work in its field of medicine. It is indispensable to those who are involved in clinical immunology or in the study of animal disease models of immunologically mediated phenomena. A consortium of expertise from both sides of the Atlantic has contributed the 63 chapters contained in this two volume work. The quality of each is uniformly high. The format of the textbook has not been altered from the first edition; obviously the editors have not wished to tamper with success. Almost 400 new pages of text, however, have been

added, thereby expanding the volume by one-third. A final chapter, an appendage, on principles of methods in immunology seems out of place in this book, and is hardly an adequate source of information on methodology for investigators in the field. It is dissonant with the highly authoritative scholarship of the preceding 62 chapters.

Almost 75 years have elapsed since Metchnikoff's publication of lectures on immunopathology. He certainly would be gratified to see how his fundamental insights into immunopathology have blossomed into a complex, vital and growing field of medicine and medical research. I think that he too would have prized this textbook.

Fred S. Rosen

Fred S. Rosen is James L. Gamble Professor of Pediatrics at Harvard Medical School and Chief of the Immunology Division, Children's Hospital Medical Center, Boston, Massachusetts.

Behaviourist history

Behaviourism and the Limits of Scientific Method. By Brian D. Mackenzie. Pp. 193 (Routledge and Kegan Paul: London and Henley, 1977.) £4.95.

In *Behaviourism and the Limits of Scientific Method* Dr Mackenzie traces the history of behaviourism from Watson to Skinner and examines its philosophical assumptions. He notes that there have been several different kinds of behaviourist. Watson's insistence that in order to be a science, psychology should be concerned only with observable behaviour and not with the furniture of the mind was followed by a more general concern with scientific method, and many behaviourists went on to embrace the then fashionable reductionist doctrines of the logical positivists and operationists. The most extreme version of behaviourism is that still being promulgated in increasing isolation by B. F. Skinner, who abjures all theoretical concepts.

Dr Mackenzie makes some interesting points. He claims for example that the learning theorists of the thirties—Hull, Tolman and Guthrie—were led astray by their adherence to logical positivism. Because they believed that their theories were merely ways of systematising the data and that their explanatory concepts had no ontological status, they failed to make explicit many of the assumptions behind their theories: the mutual misunderstandings thus engendered led to many pointless wrangles. It is certainly true that, as they later came to realise, their theories lacked rigour, but the theoretical looseness surely arose less from their philosophical assumptions than from the fact that at that time no formal language existed in which to couch rigorous explanations of behaviour.

Dr Mackenzie fails to note the difference between postulating intervening variables and putting forward a mechanism to explain behaviour. It is ironical that psychologists should have been reluctant to reify their explanatory concepts. It is clear that real processes intervene between the stimulus and the response and these processes are instantiated in a corporeal body known as the brain. To understand the logic of the brain's connectivity, psychologists needed to evolve new concepts, but the reality of the brain cannot be doubted in the same way as it is possible to doubt the reality of quarks.

Although Dr Mackenzie has provided an interesting and original account of the philosophy and history of behav-

iourism, the book has one curious fault: he has kept his eye so firmly on his quarry that he seems unaware that there have always been psychologists of many different colours. He even refers to behaviourism as "the most important single influence in the continuing development of modern psychology". This is a startling assertion to make at a time when many psychologists explicitly take account of introspection, and when almost all are concerned with building process models. Behaviourism has in fact never had much effect on the study of perception, language or cognition;

even those studying memory have shaken off its influence over the past two decades. Moreover, at the time when Hull, Tolman and Guthrie were at their most influential, Lashley, Kohler, Wertheimer and Bartlett were putting forward ideas and theories of a very different kind. Dr Mackenzie clearly did not come to praise behaviourism: it is a pity he has not appreciated that he is too late to bury it.

Stuart Sutherland

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition at the University of Sussex, UK.

Photoconductive techniques

Photoconductivity and Related Phenomena. Edited by J. Mort and D. M. Pai. Pp. 502. (Elsevier: Amsterdam and New York, 1976.) Dfl. 155; \$61.95.

PHOTOCONDUCTIVITY is probably the most sensitive of all the electrical phenomena, to the defect properties of insulators and semi-insulators. Furthermore, the photoconductive technique can be studied under a wide range of experimental variables such as wavelength, intensity, temperature, applied electric field, and in steady-state and non-steady conditions. Consequently, photoconductive data are often very distinctive and rich in structure, providing a powerful probe for investigating the trap parameters and structure of solids, as well as an insight into physical processes involved such as absorption photogeneration recombination, transport, intramolecular relaxation, electro-photon interaction, and so on.

It is not surprising, therefore, that in recent years photoconductivity techniques have been increasingly utilised in the study of the disordered and amorphous state. This book describes the applications of photoconductive techniques to these materials, a collection of review articles by leading workers in the field.

The first nine chapters are devoted to basic concepts and experimental techniques. The book opens appropriately with a review of contact effects, often neglected and ignored (conveniently?), yet absolutely fundamental to a proper understanding of photoconductivity. Chapter 2 deals with experimental techniques, with particular attention on transient and time-varying Hall and photo-electromagnetic

techniques. Chapter 3 is concerned with the time-dependent photoconductivity, an area which I believe offers the greatest scope for future study, but which poses formidable problems. Chapters 4-9 deal with photoconductivity in a variety of materials: covalent semiconductors, molecular crystals, amorphous tetrahedrally bonded solids, amorphous chalcogenides, and polymeric photoconductors. The most comprehensive of these, and a very readable account, is the one on amorphous chalcogenide and non-polar liquids. There is surprisingly little duplication considering the closely-related nature of the work, and that the works are by different sets of authors, thus reflecting the wide-ranging versatility of photoconductive techniques.

The final two chapters are concerned with applications of photoconductivity. Chapter 10 deals with photoelectronic semiconductor devices. I would have liked to have seen more space devoted to this chapter, and more emphasis on devices using disordered photoconductors, because this is what the book is really all about; even if this were at the expense of those based on crystalline materials. The last chapter, a fitting close to the book, is concerned with electrophotography, and provides a good account of the basic principles of electrophotography, image formation and development, and requirements for photoreceptor materials.

On the whole, most of the articles are concerned with underlying physical processes stressed rather than the theory, and most are very readable. The book is prepared to a high standard, and should appeal to experts (and others) in the field, although I doubt very much if the price (\$62) will.

J. G. Simmons

J. G. Simmons is Professor of Electrical Engineering at the University of Toronto, Ontario, Canada.

Simplifying sociobiology

Sociobiology and Behaviour. By D. P. Barash. (Elsevier Scientific: New York, 1977.) Hardback, \$9.95; paperback, \$4.95.

Sociobiology and Behaviour is essentially a much abbreviated version of parts of E. O. Wilson's *Sociobiology: The New Synthesis* (Harvard University Press). After introducing the basic ideas of natural selection, Barash proceeds through the by now familiar topics: altruism, living in groups, mate selection, parental investment, and aggression. Although Barash covers the 'right' topics, is usually clear in what he says, and provides plenty of good examples, this is not a book I would recommend to first-year undergraduates, the intended audience.

Although I am all in favour of making scientific texts clear and informal in style, I find Barash's racy prose annoying. Phrases such as "tunnel of love" (sticklebacks nest), "shennanigans of a love-crazed prairie chicken" (lek behaviour), "slice of the sociobiological cake" (meaning uncertain) lack both elegance and precision. This, however, is a purely aesthetic reaction, and my real objections to the book arise from three aspects of its scientific content. First, Barash claims more for sociobiology than it deserves. He proudly proposes (for the first time, he says) the Central Theorem of Sociobiology—"animals should behave so as to maximise their inclusive fitness". It sounds to me rather like what Darwin said more than a hundred years ago. Second, the book sets a bad example to students by giving the wrong impression of what scientific prediction is all about. On p259 we are told of the prediction of "sociobiologic theory" that "territories will be defended when it is adaptive for individuals to do so, i.e. when their personal inclusive fitness (*sic*) is maximised"! Almost as bad is one of the quoted 'tests' of this theory: wildebeest do not defend territories because it would not be adaptive for them to do so. It is a great pity that Barash gives readers the impression that "sociobiologic theory" (which I call natural selection) makes only tautologous or trivial and uninteresting predictions, because evolu-

tionary theory can be used to make highly quantitative and often counter-intuitive predictions just like any proper hypothesis.

My third criticism is that Barash, in his effort to make the book simple, often stops short of the point where he could convey some real understanding of the issues involved. Let me give a few examples. The discussion of costs and benefits of ritualised fighting (p233) simply misses the point because it does not make clear the fact that the benefit to an individual derived from any particular strategy depends on what everyone else does: the payoffs are frequency dependent, so the problem is to find an evolutionary stable strategy. The same line of reasoning should have been applied to the review of mating systems (p168 *et seq.*). In summarising the idea of sexual selection, Barash presents one hypothesis (the "handicap principle") without pointing out any of its known pitfalls, and fails to make a clear statement of the most widely accepted idea (Fisher's theory). The important work of Trivers and Hare on eusocial insects is incorrectly reported in a single sentence on p84; the cost of sex is incorrectly explained on p139; and although Barash asserts on p119 that "gregariousness in shorebirds represents a unique and optimal compromise between conflicting demands of predation . . . and . . . foraging interference", the reference he cites has no information on whether or not spacing is optimal.

These points, and many other similar ones, are not too difficult to explain to first-year undergraduates (see *The Selfish Gene* by Richard Dawkins, Oxford University, 1976), and I feel that it is absolutely essential to get things exactly right if one is to convince students that evolutionary biology is more than a hand-waving exercise.

A final point about presentation. The text is well illustrated with good graphs, but inexplicably the photographs often have little to do with the text. A description of distraction displays in the Killdeer is accompanied by a photograph of two Cranes, a pair of male Topis serves to illustrate territorial behaviour in the Wildebeest, and plate 3.3, a bizarre study of grey blobs, defies analysis.

John Krebs

John Krebs is Lecturer in Zoology at the Edward Grey Institute of Field Ornithology, University of Oxford, UK.

BRITISH MUSEUM (Natural History)

Recent titles in the Bulletin of the British Museum (Natural History) Zoology series

THE LARVAL AND POST-LARVAL DEVELOPMENT OF THE SCORPION SPIDER-CRAB *Inachus dorseletensis* (Pennant) (FAMILY MAJIDAE) REARED IN THE LABORATORY

R. W. Ingle

Vol.30:9

Pp.20, 10 text figs.

£2.00 post free

FREELIVING MARINE NEMATODES FROM SOUTHERN AFRICA

J. Coles

Vol.31:1

Pp.49, 14 figs, 1 plate

£4.10 post free

THE PLANKTONIC COPEPODS OF THE NORTHEASTERN ATLANTIC OCEAN: SOME TAXONOMIC OBSERVATIONS ON THE ONCAEIDAE (CYCLOPOIDEA)

G. A. Boxshall

Vol.31:3

Pp.55, 25 figs.

£4.75 post free

LARVAL DEVELOPMENT OF BRITISH PRAWNS AND SHRIMPS (CRUSTACEA: DECAPODA: NATANTIA) 1. LABORATORY METHODS AND A REVIEW OF *Palaemon* (Paleander) *elegans* Rathke 1837

A. A. Fincham

Vol.32:1

Pp.28

£2.60 post free

Many other titles on marine biology including:

THE ECHINODERMS OF SOUTHERN AFRICA

A. M. Clark

Aug. 1976

Bound 4to

Pp.vi+ 276, 276 figs.

£20.00 post free

Obtainable from booksellers, the Museum bookshop or by post (cash with order please) from:

PUBLICATIONS SALES
BRITISH MUSEUM
(NATURAL HISTORY)
CROMWELL ROAD
LONDON SW7 5BD

● The UK edition of *Neuronal Recognition*, by S. H. Barondes, has been published by Chapman and Hall (London), price £20. For review, see *Nature*, 267, 85, 1977.

Hückel molecular orbital model

The HMO Model and its Application. Vol. 2: Problems with Solutions. Vol. 3: Tables of Hückel Molecular Orbitals. By E. Heilbronner and H. Bock. Translated by W. Martin and J. Anthony Rackstraw. Pp. viii+449 and viii+190. (Wiley-Interscience: London and New York; Verlag Chemie GmbH: Weinheim, 1976.) £15 and £9; \$33 and \$18.

THESE volumes complete Heilbronner and Bock's handbook on the Hückel molecular orbital model. Volume 1 set out the basic theory and techniques, and contained numerous problems; Volume 2, which is almost as long, contains only the solutions to those problems. Anyone who works through all the problems would acquire a very extensive knowledge of Hückel theory. I suspect, however, that this knowledge would be deficient in being rather mechanical, superficial and uncritical. It is typical, for example, that although the main text provides an adequate

general derivation of the pairing theorem, the associated problem requires the reader to prove that a crude approximate formula for Hückel energies also satisfies the theorem—a result of far less power and significance. Again, the authors provide a mechanical procedure for factorising a Hückel determinant with the help of the molecular symmetry. Not only is it mechanical; it is also unnecessarily laborious, since it involves a lengthy series of operations on the entire secular determinant. Not only that; it doesn't even work where there are degeneracies (try it for cyclopropenyl). Setting up the determinant directly in terms of symmetry-adapted functions is easier, more general, and more illuminating.

Hückel theory is well known to be a very poor approximation for many purposes. The authors point out some of its failings, but do not indicate how they may be overcome, even by providing a suitable reference. Examples are the failure to predict negative spin populations, and the false degeneracy of the second excited state of alternant hydrocarbons. These are both due to the failure of Hückel theory to take account of electron repulsion; however, in both cases a simple correction can

be made to improve the Hückel result.

The third volume contains tables of Hückel molecular orbital coefficients and derived quantities such as bond orders and atom-atom polarisabilities. It covers chains (up to 10 atoms), rings (up to 18 atoms), an assortment of alternant and non-alternant molecules with up to three rings, and the partial π systems obtained by removing one carbon atom from naphthalene, anthracene, phenanthrene, azulene and biphenylene. It is intended partly as an adjunct to the problems, many of which refer to the data in the tables; however, it is also more generally useful. On the other hand, when writing a Hückel molecular orbital programme is a straightforward undergraduate exercise it cannot be said that the need for such tables is very great.

In summary, the work cannot be regarded as a definitive treatise on Hückel theory: it omits too much. As an introduction it is reasonably thorough, although even here there are some omissions. A useful but not indispensable book. **A. J. Stone**

A. J. Stone is an Assistant Director of Research in Theoretical Chemistry at the University of Cambridge, UK.

'TIBS' has become a household word in biochemistry

Trends in Biochemical Sciences

The only biochemical journal with the readability of a newspaper

Personal subscriptions
US \$ 24.00/Dfl.60.00 for 12 issues

Published for the International Union of Biochemistry by Elsevier

For a specimen copy write to:

'TIBS'

P.O. Box 548, Amsterdam
The Netherlands, or 52 Vanderbilt
Ave., New York, NY 10017

The Dutch guilder price is definitive. US \$ prices are subject to exchange rate fluctuations

Plasma in gas

Plasma Phenomena in Gas Discharges. By R. N. Franklin. (Oxford Engineering Science Series.) Pp. ix+249. (Clarendon; Oxford: Oxford University: London, 1976.) £15.

LECTURE courses spawn books as readily as the proverbial salmon spawning eggs. Unfortunately graduate courses and the books they gestate too frequently limit themselves to the particular interests of the lecturer. This would seem to be the rationale behind this scholarly but narrow book by Raoul N. Franklin.

The book is largely concerned with analysing the positive column and illustrating the diversity of plasma phenomena which can occur therein. Though these small scale, low current, experiments are of great importance in the historical development of plasma physics, they nevertheless represent a very small area in what is now a much wider experimental field.

In the first half of the book detailed solutions of radial equilibrium, of space-charge effects at walls and electrodes and of the influence of a magnetic field in a positive column are presented. The second half of the book rests a little uneasily on the first part and aims to describe the variety

of plasma phenomena that can be observed in gas discharges. It is largely, however, a catalogue of linear plasma waves, including electron plasma waves, ion waves, ionization waves, magnetosonic and 'whistler' waves and drift waves. As a description of wave phenomena in plasmas, it is a useful and reasonably comprehensive compendium, though at this level it must be compared with many excellent books on waves in plasmas. I would have been happier with more physical insight, discussion and interpretation of the equations used and results obtained, as opposed to the *ad hoc* writing down of equations, linearisation and solution of the resultant dispersion relations. As a wide catalogue and a clear analysis of waves in plasma it is useful. A helpful feature is the inclusion of comparisons with experiment whenever possible and a wide range of references to experimental work.

Though the editing can be faulted and the production is dull, the book is scholarly and precise. It is certainly a useful text for those working directly in gas discharges but of less value to the wider student of plasma physics.

David Potter

David Potter is Lecturer in Plasma Physics in the Blackett Laboratory at Imperial College, London, UK.